This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Conference on Business Cycles

Volume Author/Editor: Universities-National Bureau

Volume Publisher: NBER

Volume ISBN: 0-87014-193-7

Volume URL: http://www.nber.org/books/univ51-1

Publication Date: 1951

Chapter Title: A Test of an Econometric Model for the United States, 1921-1947

Chapter Author: Carl Christ

Chapter URL: http://www.nber.org/chapters/c4760

Chapter pages in book: (p. 35 - 130)

A TEST OF AN ECONOMETRIC MODEL FOR THE UNITED STATES, 1921-1947*

CARL CHRIST, *Cowles Commission for Research in Economics; Fellow of the Social Science Research Council*

This paper presents a revision of Lawrence Klein's sixteen-equation Model III for the United States.[1] The starting point of the revision is a test of Klein's model for 1946 and 1947 carried out by Andrew W. Marshall (17), which rejected several of Klein's equations. The equations of the revised model are estimated from a sample consisting of Klein's sample plus the two years 1946 and 1947. The estimates of the equations of the revised model are tested against the 1948 data.

In Sections 1-4 I have drawn freely and without specific acknowledgment on definitions and theorems from the published and unpublished literature, particularly on Anderson and Rubin (1), Haavelmo (5, 6, 7), and Koopmans (14, 15, 16).

I wish to acknowledge the helpful suggestions and criticism given by J. Bronfenbrenner, H. Chernoff, N. Divinsky, M. Friedman, J. Gurland, L. Hurwicz, L. Klein, T. Koopmans, H. Markowitz, J. Marschak, L. Metzler, S. Reiter, and other members of the Cowles Commission staff. I wish also to express my appreciation to the Social Science Research Council, whose fellowship aid made this study possible; to the Cowles Commission for Research in Economics, whose clerical and computing services were put at my disposal; and to the National Bureau of Economic Research, which sponsored the Conference on Business Cycles.

1 ECONOMETRIC MODELS: GENERAL CONCEPTS AND DEFINITIONS

When one thinks of science, one usually thinks also of experiments. In a typical experiment, there is *one* variable whose behavior is studied under various conditions. The experimenter fixes at will the values of all the other variables he thinks are important, and observes the one in which he is interested. He then repeats the process, fixing different values of the other variables each time. Some of these "other variables" may not be

* This paper will be reprinted as Cowles Commission Paper, New Series, No. 49.

[1] Klein (11, 13). Numbers in parentheses in contexts like this indicate references listed in Appendix G.

under his control; the important thing is that they are fixed in advance of the experiment.[2] The variables that are fixed in advance are called independent; the one that the experimenter merely observes is called dependent. The experimenter hopes to find a single equation that describes closely the relationship he has observed.

In more complicated situations there may be more than just *one* relationship among the variables studied. This is the case when there is a determinate result *and,* at the same time, there are two or more important variables that are not fixed in advance; there may even be no experimenter. Economics abounds with such situations. The simplest is of course a competitive market, in which neither price nor quantity is fixed in advance. The economist assumes that *two* relations between these variables, a supply equation and a demand equation, must be simultaneously satisfied.

The econometric work discussed here is based on the belief that we will do well to make our theory conform to this state of affairs. Accordingly we deal with systems of simultaneous equations, called structural equations.[3] Each structural equation is assumed to describe an economic relation exactly except for random shocks; hence each contains a non-observable random disturbance (with mean assumed to be zero).[4,5] Effects of errors in measuring variables are here assumed to be small relative to the disturbances (this kind of model is called a shock model, as distinct from an error model).

Time must enter into the equations if they are to describe a dynamic process. The work discussed here treats time as if it came in discrete chunks of equal size, called periods. The raw materials are time series for the variables considered. A given equation is supposed to represent a relation that holds, for any time period t, among a given set of variables evaluated as of the time period t, where t runs over a sequence of periods from 1 to T.

Note that so far we have used the term 'variable' to denote something

[2] They may be constants fixed in advance or random variables with probability distributions fixed in advance (such as weather).

[3] Structural equations are divided by Koopmans (14) into four classes: equations of economic behavior such as the consumption function, technical constraints such as the production function, institutional constraints such as tax schedules or reserve requirements, and definitional identities such as income equals consumption plus investment. Another possible type is made up of market adjustment equations, of which equilibrium conditions are a special case.

[4] This means we believe that either (1) there are systematic discoverable causes for *all* the observed variation of the variables but we are satisfied for the time being if we can explain enough of the variation so that the residual appears random, or (2) there really *are* random elements in economic affairs. For present purposes we do not care which of these two is the case.

[5] Identities are meant to be perfectly exact and hence contain no disturbances.

like national income Y or price level p, which can have different values from one period to the next. We shall continue to use 'variable' in this way, but at times we shall use it instead to denote something like national income *in period t,* denoted by Y_t, or price level *in period t,* denoted by p_t. In the first sense, Y_t and Y_{t-1} are different values of the same variable Y in different periods; in the second sense, Y_t and Y_{t-1} are different variables. It will usually be easy to tell from the context the sense in which the term is being used.

We might have G structural equations containing $G + K$ variables (second sense) at time t. Suppose that of the variables at time t, K are fixed in advance of time t, not by an experimenter to be sure, but by society, or by nature, or even by the past operation of the system of equations; they are called predetermined variables at time t. The G variables, which are not fixed in advance but are determined by the economic process we seek to describe (analogous to the dependent variable in the simple experimental case at the beginning), are called jointly dependent variables at time t (sometimes they are called current endogenous). More precisely, the definitions are as follows. (1) Variables (in the first sense) that are stochastically independent of the random disturbances in the equations are called exogenous; they may be arbitrarily fixed by some agency or process, or they may themselves be random. All other variables (in the first sense) are called endogenous. (2) Variables (in the second sense) which for time t are *either* values of exogenous variables at times t, $t - 1$, $t - 2, \cdots$, *or* lagged values of endogenous variables, i.e., values at times prior to t, are called predetermined variables at time t. Variables (in the second sense) which for time t are current values of endogenous variables, i.e., values at time t, are called jointly dependent variables at time t. Sometimes the phrase "at time t" is omitted, but whenever we speak of jointly dependent or predetermined variables, it is always understood.

Such a system of structural equations is called a *structure,* provided (1) that each equation is completely specified as to form and as to numerical values of parameters, and (2) that it is accompanied by a similarly completely specified joint probability distribution function of the disturbances.[6]

One objective will be to find a structure or structures that will enable us to rationalize past observations of economic variables and to predict future ones.[7] This will be difficult because we cannot observe structures or disturbances directly. However, we can observe samples from the joint

[6] Or, more generally, a structural relation or a distribution may be completely specified by a graph, if it is not expressible in terms of simple functional forms. But such relations are very difficult to work with.

[7] Here the "future" includes the part of the past that was not consulted in the process of finding the structure in question.

conditional probability distribution function $\phi(y \mid z)$ of the jointly dependent variables $y = (y_1, y_2, \cdots)$, given the predetermined variables $z = (z_1, z_2, \cdots)$. It is clear that any given structure generates exactly one such distribution function, i.e., to any given structure there corresponds exactly one distribution function $\phi(y \mid z)$ which is consistent with it. It is natural to ask whether this correspondence is one to one in *both* directions; in other words whether, if we knew only the conditional distribution $\phi(y \mid z)$ of the jointly dependent variables given the predetermined variables, we could proceed backwards and find a unique structure that generates it. In general the answer is no, because in general there are several (or an infinite number of) structures consistent with a given $\phi(y \mid z)$. Thus even apart from the sampling problem of estimating $\phi(y \mid z)$ from a finite sample, it is in general not possible to find a unique structure by studying observations alone. This would not be serious were it not for the fact that in general the structures generating a given $\phi(y \mid z)$ are not identical in their implications about the effects of economic policy decisions (see Sec. 2 for a more detailed discussion).

Sometimes it is possible, on theoretical grounds, i.e., on the basis of knowledge derived ultimately from other observations not used to estimate $\phi(y \mid z)$, to find a set of restrictions that we believe must be satisfied by any structure that can make good predictions. Such a set of restrictions defines a *model,* i.e., the set of exactly those structures that satisfy the given restrictions.

A model is said to be structure-identifying, or simply identifying, if each possible distribution function $\phi(y \mid z)$ is generated by exactly one structure belonging to the model. A structure is said to be identified (or identifiable) within a given model if the model contains no *other* structures generating the distribution $\phi(y \mid z)$ that is generated by the given structure. It is important to note that the problem of the identification of structures is completely separate from the problem of estimating probability distributions from finite samples, and would exist (in the context of simultaneous equations) even if there were no random elements. It is of course to be hoped that enough theoretical restrictions are available to permit the construction of identifying models.

It is desirable that the models be more general rather than less. But at this early stage of the development of econometrics it is convenient to impose further restrictions in addition to (or even at the expense of some of) those derived from theory, in order to keep the models fairly simple. For instance, it is customary for simplicity's sake (though not conceptually necessary) to choose a model that is a parametric family of structures, i.e., a set of G simultaneous equations in G jointly dependent variables and a joint distribution function of disturbances, both having a specified form but unspecified parameters.[8] Such a model is preferable because of its relative ease of handling: first, the equations and probability distribution function of the model can be set up on the basis of previous knowledge; then observations of appropriate variables can be obtained; then the parameters can be estimated by straightforward statistical procedures (see Secs. 3 and 4).[9]

Revisions of a model in the light of its performance in forecasting are of course permitted and expected; to this subject we shall return in Section 7.

To simplify computational procedures, further restrictions are placed on the model: it is assumed to be linear in the unknown parameters (though not necessarily in the variables); it deals with macro-variables (aggregates) as distinct from micro-; its disturbances are usually assumed to be normally distributed and serially uncorrelated. Obviously some of these restrictions make the model a poor approximation to the actual world. It can be expected that they will be made more realistic as statistical and economic theory, computational facilities, and data permit.

[8] Consider the model represented by the following system of equations and restrictions:

(1) $D = \alpha_1 p + \alpha_2 Y + \alpha_o + u$
(2) $S = \beta_1 p + \beta_2 w + \beta_o + v$
(3) u and v are normally distributed, with distribution function $\phi(u,v)$ and means zero.
(4) successive drawings from $\phi(u,v)$ are independent.
(5) $E(u \mid Y,w) = E(v \mid Y,w) = 0.$

This model will be changed if we restrict any of its parameters to specific values or to specific ranges, or add new terms, or change the assumptions about the distribution $\phi(u,v)$, etc. Two hypothetical structures belonging to it are:

(6) $D = -2p + .10Y + 1.5 + u$
(7) $S = 3p - 2.6w - 0.8 + v$
(8) $\sigma_u^2 = 1, \quad \sigma_v^2 = 4, \quad \sigma_{uv} = -1.5$
(9) restrictions (3) to (5) above

and

(10) $D = .16Y + 1.2 + u$
(11) $S = 2.8p - 3w - 1.3 + v$
(12) $\sigma_u^2 = 2, \quad \sigma_v^2 = 3, \quad \sigma_{uv} = 0$
(13) restrictions (3) to (5) above.

[9] In practice, the tendency is to select a model *after* looking at the data to be used to estimate its parameters. This is useful and legitimate, even necessary, as a means of *suggesting* hypotheses. However, the effect is to make spuriously small the estimated standard errors we obtain from the usual formula, i.e., to give us excessive confidence in our estimates, because this formula assumes that all the restrictions implied in our model were derived from some *a priori* source of knowledge *before* we examined the data, whereas in fact some were derived from an examination of the data.

With these restrictions, a model consists of G simultaneous equations in G endogenous variables and a distribution function of disturbances, thus:

$$\text{(A)} \quad \begin{cases} \sum_{j=1}^{J} \alpha_{gj} f^j \, (y_1', \cdots, y_G'; z_1', \cdots, z_M') = u_g, \; g = 1, \cdots, G, G \leqq J. \\ \phi(u_1, \cdots, u_G) = \text{joint normal distribution with mean zero.} \end{cases}$$

where: y_h' = jointly dependent variables at time t, $h = 1, \cdots, G$.

z_m' = predetermined variables at time t, $m = 1, \cdots, M$.

f^j = functions of the y_h' and the z_m', which are of a given form and contain *no* unknown parameters, $j = 1, \cdots, J$ (for example, X or $(pX - \mathcal{E})/q$ in Klein's model; see Sec. 5). As a special case, f^j might equal y_j' for $j = 1, \cdots, G$ and equal $z_{j\text{-}G}'$ for $j = G + 1, \cdots, J$; if so, $J = G + M$ and we have linear equations; we do not restrict the model to this extent.

α_{gj} = known or unknown parameters (some of which may be zero).

successive drawings from $\phi(u_1, \cdots, u_G)$ are independent of one another and of current and previous values of z_m', $m = 1, \cdots, M$.

each variable and disturbance is understood to carry the subscript t to indicate that it is evaluated as of period t, $t = 1, \cdots, T$.

To illustrate (A) concretely, consider the following simple income-consumption model, where C, V, and Y are respectively consumption, investment, and income:

$$\text{(A}'\text{)} \quad \begin{array}{lll} \alpha_{11}\, C/Y + \alpha_{12}\, Y & + \alpha_{15} & = u_1 \\ \qquad\qquad\quad Y - C - V & & = 0 \\ & \phi(u_1) & = N\ (0, \sigma^2) \end{array}$$

Here Y and C are the endogenous variables y_1' and y_2', and V is exogenous. $G = 2$ and $J = 5$; $f^j\,(C, Y; V)$ is equal to C/Y when $j = 1$, to Y when $j = 2$, to $-C$ when $j = 3$, to $-V$ when $j = 4$, and to 1 when $j = 5$.

$\alpha_{13} = \alpha_{14} = \alpha_{21} = \alpha_{25} = 0$; $\alpha_{22} = \alpha_{23} = \alpha_{24} = 1$; $u_2 = 0$.

N $(0, \sigma^2)$ is the normal distribution with mean zero and variance σ^2. By checking these statements, the reader can verify that (A′) is a special case of (A). Similarly, any system of equations that is linear in the unknown parameters can be expressed in the form (A).

We can rewrite (A) in a more convenient form by separating all the f^j into two classes: (1) those which involve some jointly dependent variables, i.e., some subset of the y_h' (whether or not they also involve any of the z_m'), and (2) those which are completely predetermined, i.e., involve only the z_m'. We call the first group y_i, $i = 1, \cdots, I$, and the second group z_k, $k = 1, \cdots, K$. Then the model becomes

$$(\mathrm{B}) \quad \begin{cases} \sum_{i=1}^{I} \beta_{gi} y_i + \sum_{k=1}^{K} \gamma_{gk} z_k = u_g, \; g = 1, \cdots, G. \\ \phi(u_1, \cdots, u_G) = \text{joint normal distribution with mean zero.} \end{cases}$$

where: y_i = jointly dependent variables at time t (including functions f^j which depend on any of the y_h' and have no unknown parameters), $i = 1, \cdots, I$.

z_k = predetermined variables at time t (including functions f^j which depend on no y_h' and have no unknown parameters), $k = 1, \cdots, K$.

$\beta_{gi} = \alpha_{gj}$ for all i, j such that $f^j = y_i$, $g = 1, \cdots, G$, $i = 1, \cdots, I$, $j = 1, \cdots, J$.

$\gamma_{gk} = \alpha_{gj}$ for all j, k such that $f^j = z_k$, $g = 1, \cdots, G$, $j = 1, \cdots, J$, $k = 1, \cdots, K$.

successive drawings from $\phi(u_1, \cdots, u_G)$ are independent of one another, and of current and previous values of z_k, $k = 1, \cdots, K$.

each variable and disturbance is again understood to carry the subscript t.

This is the form in which we shall use the model. In general $I \geqq G$, so that there appear to be more jointly dependent variables than equations. To complete the model, it is necessary and sufficient to include the identities that define the y_i in terms of the y_h' — thus in the case of Klein's model mentioned above, there would be an equation defining, say, y_2, thus: $y_2 = (pX - \mathcal{E})/q$. There are I such identities.

To illustrate (B) concretely, rewrite (A′) as follows, together with the identities defining the y_i:

$$(\mathrm{B}') \quad \begin{array}{lll} \beta_{11} y_1 + \beta_{12} y_2 & + \gamma_{12} z_2 & = u_1 \\ \qquad y_2 + y_3 + z_1 & & = 0 \\ & y_1 & = C/Y \\ & y_2 & = Y \\ & y_3 & = -C \\ & \phi\,(u_1) & = N\,(0, \sigma^2) \end{array}$$

Here $G = 2$ again of course, $I = 3$, and $K = 2$; $z_1 = -V$, and $z_2 = 1$. $\beta_{11} = \alpha_{11}$, $\beta_{12} = \alpha_{12}$, $\beta_{13} = \alpha_{13} = 0$, $\beta_{21} = \alpha_{21} = 0$, $\beta_{22} = \alpha_{22} = 1$, $\beta_{23} = \alpha_{23} = 1$, $\gamma_{11} = \alpha_{14} = 0$, $\gamma_{12} = \alpha_{15}$, $\gamma_{21} = \alpha_{24} = 1$, $\gamma_{22} = \alpha_{25} = 0$. The reader can verify that (B′) is a special case of (B), and that any set of equations of the form (A) can be rewritten in the form (B).

Simple procedures are available for determining whether a model like

(B) is structure-identifying, i.e., whether the structures belonging to (B) are identified. A structure is said to be identified within the model (B) if and only if all its equations are. A necessary condition for the identification of an equation of (B) (if the covariance matrix of the disturbances at time t is completely unknown) is the order condition: K^{**}, the number of predetermined variables z_k in the other equations of the model but *not* in that equation, must be greater than or equal to $H - 1$, where H is the number of jointly dependent variables y_i in that equation. (If K^{**} is greater than $H - 1$, the equation is sometimes said to be overidentified, and to have $K^{**} - H + 1$ overidentifying restrictions.) There is a necessary and *sufficient* condition as well (the rank condition); it is more difficult to apply to an equation in a system where there are unknown parameters, but if it is not satisfied we expect to be notified of the fact when we reach the estimation stage by the presence of large estimated sampling variances of our estimates of the parameters (Koopmans, 15, 16).

There remains one more general remark, concerning the generation of cyclical patterns by an econometric model of the type defined here. Such a model contains lagged values of many of its variables, and therefore is a set of simultaneous difference equations. Whereas the solution of a set of ordinary simultaneous equations (given the values of the exogenous variables) is simply a set of numbers, each giving a single value for one endogenous variable, the solution of a set of difference equations (given the values of the exogenous variables) is a set of functions of time, each giving a path in time for one endogenous variable. Such a time-path gives the future history that an endogenous variable of the model would have, as a function of future values of exogenous variables, if future disturbances were zero. This history may behave in one of several ways as t increases indefinitely:

1) approach a finite limit monotonically.
2) approach a finite limit with oscillations of diminishing amplitude.
3) oscillate indefinitely with constant finite amplitude.
4) approach infinity (positive or negative) monotonically.
5) approach infinity (positive and/or negative) with oscillations.[10]

The oscillations, if any, will have a constant period and a constant 'damping ratio' (which may be less than, equal to, or greater than one according as 2, 3, or 5 above is the case) *if* future disturbances are zero and future exogenous variables are constant. Their limits as t increases indefinitely can be computed. Derivatives of period, damping ratio, and

[10] Except during periods of hyperinflation, etc., we expect the solutions of the equations we construct for our economy to be of the first, second, or third kind, i.e., not to 'explode'. This expectation is not included among the restrictions used in the estimation procedures of this paper, however.

amplitude with respect to exogenous variables and parameters can also be computed.[11]

Since econometric models can thus generate cyclical fluctuations that respond to changes in exogenous variables, they recommend themselves as promising analytical tools for business cycle research.[12]

2 PREDICTION: THE REDUCED FORM OF A SYSTEM OF STRUCTURAL EQUATIONS

The ultimate test of an econometric model, as of any theory, comes with checking its predictions. A model of the form of equations (B) in Section 1 is not ready to make predictions of the jointly dependent variables, even if its parameters are known. It must first be solved algebraically, so that each jointly dependent variable is expressed in terms of things that can be at least approximately known when the predictions are being made, i.e., parameters and predetermined variables. If (A) is linear in the y_h', then $y_h' = y_i$ for $h = i$, and (A) and (B) are just alike. Then the solution of (B) or (A) for the jointly dependent variables is simply the solution of a set of G linear equations for G unknowns y_i. It is called the *reduced form* of the model, and looks like this:

$$
\text{(C)} \quad \begin{cases} y_i = \sum_{k=1}^{K} \pi_{ik} z_k + v_i, \; i = 1, \cdots, I. \\ \phi(v_1, \cdots, v_I) = \text{joint distribution function of } v_1, \cdots, v_I. \end{cases}
$$

where: π_{ik} = parameters dependent upon the structural parameters β_{gi} and γ_{gk}, $g = 1, \cdots, G$, $i = 1, \cdots, I$, $k = 1, \cdots, K$.

v_i = a random disturbance, equal to a function of the β_{gi} and u_g, $g = 1, \cdots, G$, $i = 1, \cdots, I$.

all other symbols have the meaning given in Section 1; in particular the y_i are jointly dependent variables at time t (including functions thereof with no unknown parameters), and the z_k are predetermined variables at time t (including functions thereof with no unknown parameters), as defined in Section 1.

We still call (C) the reduced form of (B), even if (A) is not linear in the y_h', though in that case (C) is not actually the solution of (B), but is only a kind of linear approximation of it. The exact solution of (B) for either the y_h' or the y_i is nonlinear in the z_k if (A) is not linear in the y_h'. (C) is a solution of (B) in the sense that each equation of (B) is a linear combination of the equations of (C), as the reader may verify. Since the linear form (C) is more convenient for later purposes than the exact solu-

[11] This applies to nonoscillatory solutions too, except for the period.

[12] See, for example, Frisch (4), Kalecki (10), and Tinbergen (23).

tion of (B), we shall use it, and unless otherwise specified, the term reduced form hereafter will refer to it.

We have encountered the reduced form in Section 1, but not under its present name: it specifies the conditional probability distribution $\phi(y \mid z)$ of the jointly dependent variables y given the predetermined variables z.

If the parameters of the reduced form (C) are known or estimated, predictions of any desired jointly dependent y_i for time t may be made simply by substituting the values of the predetermined z_k for time t into the reduced-form equation for the desired y_i.[13] (It is also possible to predict from the structural equations, by the following process: first substitute into the structural equations the known or assumed values of the predetermined variables, and the estimated values of the structural parameters; then solve the system simultaneously to get the predicted values of the jointly dependent variables. This method in effect uses the exact reduced form rather than the linear approximation for predicting; it gives predictions identical with those of the linearized reduced form if and only if the model is linear and just-identifying; in other cases it presumably gives better predictions because it ignores fewer of the available *a priori* restrictions, but in nonlinear models it is more difficult to apply.) As we shall find in Sections 3 and 4, it is easier to obtain suitable estimates for the parameters of the reduced form (C) than for the parameters of the structural equations (B). Thus it appears that in order to predict the jointly dependent variables y_i, we need only know the values of the parameters of the reduced form, not bothering with the structural equations at all. This is true if, between the period for which the reduced form equations are estimated and the period for which predictions are to be made, no change occurs in the distribution $\phi(y \mid z)$, which means no change occurs in the structure that generates the distribution $\phi(y \mid z)$. But if the structure does change, $\phi(y \mid z)$ and hence the reduced-form parameters π_{ik} will change also, and this will invalidate any predictions based upon knowledge of the old π_{ik}. To be valid, predictions must be based upon knowledge of the new π_{ik}. To obtain this we must have (besides knowledge of the old π_{ik}) knowledge of the structural parameters β_{gi} and γ_{gk} before the structural change, and knowledge of the effect upon them of the structural change. Thus, for prediction under known structural change, the reduced form is not enough; we must know the structural parameters as well.[14]

[13] Strictly, a prediction of this kind specifies a probability distribution, not a number. Loosely, we shall use the terms 'predicted value' or 'prediction' to mean 'expected value of predicted distribution'.

[14] Even if the change in a structural parameter is known in *direction* only, not in magnitude, it is still true except in special cases that to find the *direction* of the resulting change in some variable (such as national income) it is necessary to know the *magnitudes* of other parameters of the system. See Samuelson (22), pp. 12-14.

One of the most interesting uses of econometric models will, I hope, be to predict the consequences of alternative public policy measures in order that enlightened decisions can be made. Institution of a new policy can often be interpreted as a change of structure. Information about the effect of such a structural change is likely to be available. Thus prediction under structural change can be expected to assume a very important role.

3 STATISTICAL ESTIMATION PROCEDURES: THE STRUCTURAL PARAMETERS

As we have said, we hope to be able to construct a structure-identifying model, i.e., a model containing exactly one structure able to generate each distribution function $\phi(y \mid z)$ of the jointly dependent variables given the predetermined variables. It is the job of a statistical estimation procedure, given the observations, to find estimates of that structure. We mention three such procedures: the full information maximum likelihood method (which we shall call full information for short), the limited information single-equation maximum likelihood method (which we shall call limited information for short),[15] and the least squares method.

Before describing these three estimation procedures, we shall briefly discuss maximum likelihood estimates in general. Given a sample of observations, the maximum likelihood estimate $\hat{\theta}$ of a parameter θ of the distribution of the observations is the value of θ among all possible values of θ, that yields the highest probability density for the given sample of observations. It is obtained by two steps: (a) forming the *likelihood function* of the unknown parameters given the sample of observations (this is the probability density function of the observable variables, with the actual observed values of the variables substituted into it, so that it is a function only of parameters); (b) maximizing this likelihood function with respect to the unknown parameters, treating them as variables for the moment. For a wide class of distribution functions (including asymptotically normal distributions), maximum likelihood estimates are asymptotically normally distributed. Furthermore, under certain assumptions, they are consistent,[16] and efficient[16] compared with any other estimates that are both consistent and asymptotically normally distributed.[17]

[15] The limited information method is treated in Anderson and Rubin (1).

[16] An estimate t of a parameter θ is said to be consistent if $\underset{N \to \infty}{Lim}\, Pb(\mid t - \theta \mid > \mathcal{E}) = 0$ for any $\mathcal{E} > 0$, where N is the sample size and $Pb\ (x)$ means the probability of x occurring. A consistent estimate t is said to be efficient compared with another consistent estimate t' if $\underset{N \to \infty}{Lim}\, N{\cdot}E(t - \theta)^2 \leq \underset{N \to \infty}{Lim}\, N{\cdot}E(t' - \theta)^2$.

[17] The proof that maximum likelihood estimates have these optimal properties is based on the assumption that there is a true structure, which belongs to the model used. Hence the optimal properties of maximum likelihood estimates may not exist if an unrealistic model is used. This is mentioned here as a caution, even though it is

1) The full information maximum likelihood method (full information for short) treats all the jointly dependent variables alike, considering them all dependent as a group on the predetermined variables. It consists of two steps: (a) forming the joint likelihood function of all the parameters, given the observations of all the jointly dependent and predetermined variables in the model; (b) maximizing this likelihood function with respect to all the parameters simultaneously, subject to all the restrictions implied in the model. Its application therefore requires complete specification of the model, plus observed values for all the variables included in the model. Estimates so derived are consistent, and efficient compared to any other estimates which are both consistent and asymptotically normally distributed.

2) The limited information single-equation maximum likelihood method (limited information for short) treats each equation of the model separately, not the whole model at once, but nevertheless recognizes the simultaneous-equations character of the model. For any given equation, it consists of two steps: (a) forming the joint likelihood function of the parameters in the equation, given the observations of the jointly dependent variables in the equation and all the predetermined variables in the model; (b) maximizing this likelihood function with respect to the parameters in the given equation, subject to the restrictions implied in the probability distribution function of the disturbances and in the given equation (including in particular the restrictions stating which predetermined variables do not enter the given equation).[18] Its application to a particular equation therefore uses complete specification of that equation, plus observed values for the jointly dependent variables appearing in that equation, plus observed values for all the predetermined variables appearing in the model. Estimates so derived are consistent, and efficient compared with any other

a commonplace that the results of any statistical analysis may fall down if the original assumptions are not fulfilled.

The proof is based also on the assumption of normally distributed disturbances. If this assumption is not true, estimates computed *as if* it were true still retain the property of consistency; such estimates are called quasi-maximum likelihood estimates.

Proofs of consistency in estimation employ also the assumption that the matrix of moments of predetermined variables is bounded in the limit as the sample size increases.

[18] In the process of maximizing this likelihood function, what is done essentially is first to compute the least squares estimates of the parameters of the reduced form (see Sec. 4), and second to transform those estimates into estimates of the structural parameters by means of the inverse of the transformation used to obtain the reduced form (C) from the structural equations (B). This is a complex process only in the case of *overidentified* structural equations.

estimates that (a) are consistent, (b) are asymptotically normally distributed, *and* (c) use the same information, i.e., the same *a priori* restrictions and the same observations. The limited information estimates are not efficient compared with the full information estimates (in cases where there is any difference), because the latter use more restrictions and more observations.

In another variant of the limited information method the likelihood function that is maximized is an abbreviated version of the one assumed to be the true one: it is conditional not upon all the predetermined variables appearing in the model, but only on a subset of them. The subset must include all the predetermined variables appearing in the equation to be estimated, and at least $H - 1$ others (where H is the number of jointly dependent variables in the equation to be estimated), but otherwise it is arbitrary. In other words, this subset must be large enough to ensure that the given equation satisfies the necessary condition for identification stated in Section 1. This abbreviated variant of the limited information method uses observations on only the jointly dependent and predetermined variables in the given equation plus the $H - 1$ (or more) other predetermined variables included in the likelihood function. Thus for overidentified equations it yields estimates that (for finite samples) are not unique because the choice of the $H - 1$ or more other predetermined variables is arbitrary. Estimates of structural parameters obtained by the abbreviated variant of the limited information method are consistent. They are less efficient than ordinary limited information estimates (because the latter use more observations and correct instead of incorrect restrictions), but they are efficient compared with any other consistent and asymptotically normal estimates that use the *same* observations and restrictions.[19]

3) The least squares method treats each equation of the model completely separately, as if there were no other equations. It is not a maximum likelihood method except in special cases. For any given equation, it consists of three steps: (a) choosing arbitrarily one variable to be regarded as dependent upon the others; (b) forming the likelihood function of the parameters in the equation, given the observed values of the dependent variable and the other (independent) variables in the equation; (c) maximizing this likelihood function with respect to the parameters in the equation. Its application to a particular equation requires complete specification of the equation, plus observed values for only those variables appearing in the equation. In the general case of a model with more than one equation,

[19] There are still other variants, not discussed here, called limited information subsystem maximum likelihood methods, in which a proper subset of two or more equations of the structure is estimated simultaneously; see Rubin (20).

least squares estimates of the parameters are biased and inconsistent, as Haavelmo (7) has proved. They are also arbitrary within certain limits because of the arbitrary choice of a dependent variable.

In certain cases, depending upon the restrictions implied in the model, the full information method is equivalent to the least squares method, and all three methods lead to identical computation procedures and estimates.[20]

The full information method is most expensive, and therefore has never (to my knowledge) been used for a system of more than three equations, except where for simplicity all disturbances at time t were assumed to be independent of one another. The limited information method, though less efficient, is considerably less expensive and has been more extensively used. The least squares method, though known to give biased estimates except in special cases, is computationally much the simplest, and has been traditionally used.

The justifications advanced for the continued use of the least squares method in cases where it yields biased estimates are of two kinds. First, assuming that estimates having the asymptotic properties of consistency and efficiency *are* in fact superior (in terms of the expected value of the square of the error) for small samples as well as for large (although this superiority is proved only for large), the cost of superior estimation may be too high. Second, in the interesting cases the least squares method's bias may be small and the convergence of its estimates as the sample size increases may be rapid, so that its expected errors may be smaller for small samples than those of the full or limited information maximum likelihood methods.[21]

[20] An example is the following model, due to Hurwicz:

$$y_1 + \beta y_2 \qquad = u_1$$
$$y_2 + \gamma z = u_2$$

u_1 and u_2 are normally distributed, and serially independent

$$Eu_1 = Eu_1u_2 = Eu_2 = 0$$
$$E(u_1 \mid z) = E(u_2 \mid z) = 0.$$

The full information estimate of β from a sample of T is

$$\hat{\beta} = -\frac{\sum_{t=1}^{T} y_1 y_2}{\sum_{t=1}^{T} y_2^2}$$

which is also the limited information estimate and the least squares estimate. Note that y_2 is exogenous to the first equation of the model because its distribution depends only on z and u_2, so that it is independent of u_1.

[21] Hurwicz has pointed out informally that when estimates of some parameters are obtained using *incorrect* assumptions about the values of others, there is in general both a gain and a loss in accuracy of estimation (measured by the expected value of the squared difference between the estimate and the parameter), as compared with

The least-squares method and the abbreviated variant of the limited information method will be used in this paper. Of these two, the latter appears preferable because it preserves the simultaneous-equations character of economic theory, rather than distorting it by forcing it into a framework designed for only one dependent variable. But until we can recognize in advance the cases in which least-squares error, i.e., bias plus sampling error, in small samples is in fact so large as to make the least-squares method poorer for small samples than the limited information method, it may be just as well to use both methods and compare their results.

4 STATISTICAL ESTIMATION PROCEDURES: THE PARAMETERS OF THE REDUCED FORM

Since the equations of the reduced form contain only one dependent variable each, they are automatically identified. We mention two procedures for estimating the parameters of the reduced form. One is the ordinary least-squares method, and the other is a modification of it which for the purposes of this paper we shall call the restricted least squares method.

1) The ordinary least-squares method is equivalent to forming for each reduced-form equation the likelihood function of its parameters, given the observed values of its dependent variable and its predetermined variables, and then maximizing this likelihood function with respect to the parameters of the equation. Ordinary least-squares estimates of the parameters of any reduced-form equation are unbiased and consistent, provided either that none of the predetermined variables in the model is excluded from the reduced-form equation or that no excluded predetermined variable is correlated with any of the other predetermined variables. But the resulting estimate of the expected value of the jointly dependent variable will remain unbiased and consistent even if some of the predetermined variables are excluded and are correlated with other included ones.[22]

2) The restricted least squares method is the same as the ordinary least-squares method except that for a given reduced-form equation the maxi-

---

the results obtained if the incorrect assumptions are simply dropped. The gain comes from the faster convergence of the estimates to their expected values, and becomes smaller as the sample grows. The loss comes from the fact that the expected values to which the estimates converge are not the true values of the parameters, and is not affected by the sample size. In small samples, therefore, the gain may exceed the loss. Thus least-squares estimates, even if based on incorrect assumptions about the covariances of the disturbances, may be superior in some small sample cases to consistent methods.

[22] Furthermore, the limited information estimates of structural parameters, obtained from least-squares estimates of the parameters of the reduced form as indicated in note 18, will remain consistent as well.

mization is performed subject to a restriction or restrictions implied in the form of a (proper or improper) subset of the set of all those structural equations that contain the jointly dependent variable appearing in the given reduced-form equation.[23] The procedure yields estimates that (for finite samples) are not unique because of the arbitrary choice of the subset of structural equations. If limited information single-equation estimates of the structure have already been computed, it is much the simplest to choose a one-element subset of structural equations because most of the computations have already been made (this is the procedure followed in this paper). The restricted least-squares estimates of the parameters of the reduced form are unbiased and consistent, with the same qualifications as apply to the ordinary least-squares estimates (see the preceding paragraph). They can be expected to be more efficient than the ordinary least-squares estimates because they use more restrictions.

5 KLEIN'S MODEL III

Klein's model III has 15 equations, of which 3 are definitional identities containing no disturbances and no unknown parameters. Thus there are 12 stochastic equations to be estimated.

There are 15 endogenous variables (in the sense of the y_g' in equations (A) in Sec. 1):

C = consumer expenditures, in billions of 1934 dollars.
D_1 = gross construction expenditure for owner-occupied one-family nonfarm housing, in billions of 1934 dollars.
D_2 = gross construction expenditure for rented nonfarm housing, in billions of 1934 dollars.
H = inventories at year end, in billions of 1934 dollars.
I = net private producers' investment in plant and equipment, in billions of 1934 dollars.
i = average corporate bond yield, in per cent.
K = stock of private producers' fixed capital at year end, in billions of 1934 dollars.
M_1^D = active cash balances = demand deposits + currency outside banks, in billions of current dollars.
M_2^D = idle cash balances = time deposits, in billions of current dollars.
p = general price level, 1934:1.0.
r = nonfarm rent index, 1934:1.0.
v = fraction of nonfarm housing units occupied at year end, in per cent.
W_1 = private wages and salaries, in billions of current dollars.
X = private output (except housing services), in billions of 1934 dollars.
Y = disposable income, in billions of 1934 dollars.

[23] Appendix E gives a fuller description of the nature of these restrictions and how they are applied.

Fourteen variables are assumed to be exogenous (in the sense of the exogenous z_k' in equation (A) in Sec. 1):

D_3 = gross construction expenditures for farm housing, in billions of 1934 dollars.
D'' = depreciation on all housing, in billions of 1934 dollars.
$\mathcal{E}$ = excise taxes, in billions of current dollars.
$\mathcal{E}_R$ = excess bank reserves, in millions of current dollars.
ΔF = increase in number of nonfarm families, in thousands.
G = government expenditures (except transfers and net government interest) + net exports + net investment of nonprofit institutions, in billions of 1934 dollars.
N^S = nonfarm housing units at year end, in millions.
q = price index of capital goods, 1934:1.0.
q_1 = construction cost index, 1934:1.0.
R_1 = nonfarm housing rents, paid or imputed, in billions of current dollars.
R_2 = farm housing rents, paid or imputed, in billions of current dollars.
T = government revenues − net government interest − transfers + corporate saving,
= net national product − disposal income, in billions of 1934 dollars.
t = time, in years; $t = 0$ in 1931.
W_2 = government wages and salaries, in billions of current dollars.

Data for these variables for 1921-41 are presented in Klein (11, 13) and in Appendix A of this paper.

The 12 equations and 3 identities are as follows (they are here grouped as they were by Klein for his limited information estimation, and renumbered by me):

(1) demand for investment
$$I = \beta_0 + \beta_1 \frac{pX - \mathcal{E}}{q} + \beta_2 \left(\frac{pX - \mathcal{E}}{q}\right)_{-1} + \beta_3 K_{-1} + u_2$$

(2) demand for inventory
$$H = \gamma_0 + \gamma_1 (X - \Delta H) + \gamma_2 p + \gamma_3 H_{-1} + u_3$$

(3) output adjustment
$$\Delta X = \mu_0 + \mu_1 (u_3)_{-1} + \mu_2 \Delta p + u_{12}$$

(4) demand for labor
$$W_1 = a_0 + a_1 (pX - \mathcal{E}) + a_2 (pX - \mathcal{E})_{-1} + a_3 t + u_1$$

(6) demand for consumer goods $C = \delta_0 + \delta_1 Y + \delta_2 t + u_4$

(7) demand for owned housing $D_1 = \varepsilon_0 + \varepsilon_1 \frac{r}{q_1} + \varepsilon_2(Y + Y_{-1} + Y_{-2}) + \varepsilon_3 \Delta F + u_5$

(8) demand for dwelling space $v = \eta_0 + \eta_1 r + \eta_2 Y + \eta_3 t + \eta_4 N^S + u_7$

(9) rent adjustment $\Delta r = \theta_0 + \theta_1 v_{-1} + \theta_2 Y + \theta_3 \frac{1}{r_{-1}} + u_8$

(10) demand for rental housing $D_2 = \zeta_0 + \zeta_1 r_{-1} + \zeta_2 (q_1)_{-1} + \zeta_3 (q_1)_{-2} + \zeta_4 i + \zeta_5 \Delta F_{-1} + u_6$

(15) demand for active dollars $M_1{}^D = \iota_0 + \iota_1 p(Y + T) + \iota_2 t + \iota_3 p(Y + T)t + u_9$

(16) demand for idle dollars $M_2{}^D = \kappa_0 + \kappa_1 i + \kappa_2 i_{-1} + \kappa_3 (M_2{}^D)_{-1} + \kappa_4 t + u_{10}$

(11) interest adjustment $\Delta i = \lambda_0 + \lambda_1 \mathcal{E}_R + \lambda_2 i_{-1} + \lambda_3 t + u_{11}$

(12) definition of net national product $Y + T = C + I + \Delta H + D_1 + D_2 + D_3 - D'' + G$

(13) definition of X $X = Y + T - \frac{1}{p} (W_2 + R_1 + R_2)$

(14) definition of K $\Delta K = I$

Observe in Klein's model the following nonlinear functions involving no unknown parameters: $(pX - \mathcal{E})/q$, $pX - \mathcal{E}$, r/q_1, $1/r_{-1}$, $p(Y + T)$, $p(Y + T)t$, and $(1/p)(W_2 + R_1 + R_2)$.

Equations 1, 2, 3, and 6 are related to the market for goods and services, excluding labor and the construction of housing (these two markets will be treated separately immediately below). Demand for consumer goods (6) is a linear function of income and trend.

Demand for net investment in plant and equipment (1) is a linear function of (a) present and lagged values of deflated (by capital goods prices) privately produced national income at factor cost excluding housing, which is similar to profits, and of (b) the stock of plant and equipment at the beginning of the year. This function is meant to show the dependence of demand for investment upon (a) anticipated profits and (b) existing capital.

Demand for inventory stocks to hold (2) is a linear function of sales, of expected price change (assumed to be given by a linear combination of current and lagged prices), and of the stock of inventories at the end of the year (an inertia factor). Lagged prices do not appear because Klein found them to be unimportant statistically.

Equation 3 expresses the change in private nonhousing output as a linear function of unintended inventory accumulation (assumed to be measured by $(u_3)_{-1}$, the lagged disturbance in the demand-for-inventory-stocks equation 4), and of the rate of change in general prices. It is essentially a supply equation.

Equation 4 gives the demand for labor, measured by the total wage bill, as a linear function of trend, and of current and lagged values of privately produced national income at factor cost excluding housing (which is supposed to reflect anticipated receipts from sales, net of excises). Observe that this equation could be omitted without impairing the completeness of the model, because the variable W_1 (wage-bill) does not appear in any other equation; in other words, if this equation were omitted, a system of 14 equations in 14 variables would remain.

Equations 7-11 pertain to the housing market. Demand for owner-occupied one-family nonfarm housing construction (7), which is purchased by consumers, is a linear function of the real value of rents (where the deflator is construction costs), of accumulated cash balances (assumed to be proportional to the sum of incomes during the 3 most recent years), and of the increase in the number of nonfarm families.

Demand for rented nonfarm housing construction (10), which is purchased by entrepreneurs, is a linear function of lagged rents, of anticipated prices of housing (assumed to be given by a linear combination of construction costs lagged one and two years), of corporate bond yield, and of lagged increase in the number of nonfarm families.

Equation 9 describes the determination of the nonfarm rent level, which occurs in the housing-construction equations 7 and 10, by a linear function of lagged rents, lagged occupancy rate, and income.

Equation 11 describes the change in corporate bond yield, which occurs in the rental housing construction equation 10 and in the idle balances equation 16, as a linear function of excess reserves, of lagged interest rate, and of trend. Note that it has only one dependent variable.

Nonfarm occupancy rate, v (8), which occurs lagged in the rent adjustment equation 9, is a linear function of rents, of income, of trend, and of the supply of nonfarm dwelling units. It could, like W_1, be dropped together with its equation 8, since v occurs nowhere else in the model except in lagged form.

Equations 12-14 are definitions containing no disturbances. Equation 12 is an identity defining net national product as a sum of demand for consumer goods, net investment, increase in inventories, housing construction (net), and goods for government use. This sum might be regarded as an aggregate demand; the fact that it is called the definition of net national product indicates that implicit in the model is an assumption that quantity supplied always equates itself to quantity demanded, except for unintended inventory; see equation 3.

Equation 13 defines privately produced real output excluding housing services, which appears in equations 1-4.

Equation 14 defines stock of capital, which appears lagged in the demand for investment equation.

Klein included in his model an equation defining R_1, nonfarm rent, which he classified as endogenous. R_1 is actually exogenous, however, according to the way he treats it,[24] and we shall so regard it.

Equations 15 and 16 could, like 4 and 8, be omitted without impairing the completeness of the model, since the variables $M_1{}^D$ and $M_2{}^D$ (active and idle balances, respectively) occur in no other equations. Demand for active balances (15) is a nonlinear function of disposable money income and trend. Demand for idle balances (16) is a linear function of current and lagged corporate bond yield, of lagged idle balances, and of trend.

Equations 1, 2, 4, 6, 7, 8, 10, 15, and 16 are demand equations, describing the behavior of various economic groups in the population. Equations 3, 9, and 11 are market adjustment equations describing responses of certain market variables to disequilibria. Equations 12-14 are identities describing definitional relationships.

Klein's estimates of the parameters of his model, for both least squares and limited information methods, appear in Section 10 below.

The results I am interested in presenting are those flowing from my revision of Klein's model. This revision is based upon a test of Klein's model carried out by Andrew W. Marshall. The next section discusses Marshall's test and its findings.

[24] Its time series is obtained as shown in Appendix B and in Klein (13); its defining equation is not used at all except to obtain estimates of its value for 1919-20, for which data are lacking.

6 MARSHALL'S TEST OF KLEIN'S MODEL III

Marshall (17) tested Klein's model III, together with Klein's limited information estimates of its structural parameters, in two ways. Both ways use the *calculated* disturbances to Klein's structural equations for 1946 and 1947. For any time period t, these calculated disturbances (called u_t^*) are obtained from the structural equations by substituting into them the limited information estimates of the structural parameters, together with the values of all the jointly dependent and predetermined variables at time t.

1) Marshall's first test examines each u^* for 1946 and for 1947 to see whether it is larger than would be expected under the hypothesis that Klein's model and estimates describe 1946 and 1947 as well as they describe the sample period. This is done for each structural equation separately by means of a tolerance interval[25] for the calculated disturbances u^*: the hypothesis is accepted for a given equation and a given post-sample year if the value of u^* for that equation and that year falls inside its tolerance interval. Marshall chooses $\gamma = 0.99$ and $P = 0.99$, which means that under the hypothesis the probability is 0.99 that the tolerance interval for a given equation will include at least 0.99 of the population of calculated disturbances u^*.

A tolerance interval is of the form $\bar{x} \pm ks$, where $\bar{x}$ and s are the mean and standard deviation computed from a sample of N, and k is a number depending upon γ, P, and N.[26] In this case, $\bar{x}$ is $\bar{u}^*$, which is zero by the construction of the estimates of the structural parameters. For each structural equation, Marshall uses in place of s an estimated approximation to the standard deviation of the calculated disturbance u^*, analogous to the Hotelling (9) formula for the standard error of forecast from a regression. This approximation, which we call σ^*, is given by Rubin (21). For the g^{th} structural equation and the year t, it looks like this:

$$(17)\quad \sigma^{*2}(g,t) = E(u^{*2})$$
$$= \sigma^2 + \frac{\sigma^2}{T} + tr\Lambda\Omega + \frac{\sigma^2}{T} z_t^* M^{-1}{}_{z^*z^*} z_t^{*\prime} + z_t^o \Pi^{**\prime} \Lambda \Pi^{**} z_t^{o\prime}$$

[25] A tolerance interval is a random variable; it is an interval that encloses, with a certain probability γ, at least a certain proportion P of the individuals in a given probability distribution. This, and not a confidence interval, is what we want here: we are interested in predicting a *future drawing* from our population of years, not in the true mean. Tolerance limits for the normal distribution have been developed by Wald and Wolfowitz (24), and tables have been prepared for constructing them; see Eisenhart, Hastay, and Wallis (3). The size of the tolerance interval depends upon an estimate of the variance of the *calculated* disturbances in the sample period, i.e., it depends partly upon the estimates of the parameters of the equation.

[26] See table in Eisenhart, Hastay, and Wallis (3), pp. 102-7.

where: $\sigma^2 = E(u^2)$

T = number of years in sample.

Λ = covariance matrix of the estimates of parameters of those endogenous variables y_i appearing in the g^{th} structural equation.

Ω = covariance matrix of disturbances v_i of reduced-form equations containing those endogenous variables y_i appearing in the g^{th} structural equation.

z_t^* = vector of values in year t of all those predetermined variables z_k appearing in the g^{th} structural equation, measured from their sample means.

$M_{z^*z^*}$ = moment matrix of z^* with z^*; $m_{ij} \doteq \frac{1}{T}\Sigma(z_i - \bar{z}_i)(z_j - \bar{z}_j)$

z_t^o = vector of residuals at time t of the regressions of the z^{**} (i.e., predetermined variables z_k appearing in the system as a whole but not in the g^{th} structural equation, measured from the sample mean) on the z^*; i.e., $z_t^o = z_t^{**} - M_{z^{**}z^*} M^{-1}_{z^*z^*} z_t^*$.

Π^{**} = matrix of reduced-form parameters of the z^{**} in those reduced-form equations containing those endogenous variables y_i appearing in the g^{th} structural equation.

For each structural equation the values of T, z_t^*, and z_t^{**} are known, and estimates are available for σ^2, Λ, Ω, $M_{z^*z^*}$, $M_{z^{**}z^*}$, and Π^{**}.[27] Thus an estimate of σ^* is available for each structural equation. We call this estimate s^*.

The test for year t then takes the form of constructing a tolerance interval, $\pm ks^*$, for each structural equation, and rejecting the equation if its calculated disturbance u^* falls outside the interval. I shall call it the structural equation tolerance interval test, provisionally, or the SETI test for short.[28]

In applying the SETI test, Marshall computed ks^* in five steps, ks_1^*, ks_2^*, ks_3^*, ks_4^*, and $ks_5^* = ks^*$, corresponding to the first term of the estimate of 17, the first two terms, $\cdots$, and all five terms. For each equation he compared each of these successively with u^*, and stopped as soon as he got a region $\pm ks_i^*$ which enclosed u^*. In this way he saved some computational effort, because he did not have to compute all the terms of s^* for every equation.

2) Marshall's second test examines each calculated disturbance u^* for 1946 and 1947 to see whether it is larger than the error one would expect to make by using what he calls "naive models". Naive model I says that next year's value of any variable will equal this year's value plus a random normal disturbance; naive model II says it will equal this year's value plus

[27] See Anderson and Rubin (1).

[28] The name was suggested by John Gurland.

the change from last year to this year plus a random normal disturbance.[29]

For each naive model and each Klein structural equation, Marshall compares the calculated disturbances u^* of 1946 and 1947 with a *tolerance interval* for the calculated disturbance of the one naive-model equation that contains the variable appearing on the left side of the given Klein equation. If both u^*'s for a given Klein equation are outside the interval, Marshall rejects the Klein equation; if one is outside, he puts the Klein equation on probation; if neither is outside, he accepts the Klein equation.

From the viewpoint of this paper, the naive model tests should be applied to the calculated disturbances of the *reduced form,* not to those of the structural equations: the naive model tests are best suited to compare different methods of predicting (because their disturbances are their errors of prediction), and the predictions made by an econometric model are obtained from its reduced form (see Sec. 2), not directly from its structural equations.[30] But if a naive model test *is* applied to the calculated disturbances of the reduced form, and if it is to be a fair comparison between methods of prediction, then the treatment of the errors of the naive model should be symmetrical with the treatment of the errors, i.e., the calculated disturbances, of the reduced form of the econometric model. This means that a direct comparison of errors should be used, instead of a tolerance interval procedure such as Marshall's which will not reject an equation of the reduced form of the econometric model unless the latter's errors are about three times as large as the naive model's errors (because Marshall's value of k in his naive model tests is about 3).

The results of Marshall's SETI test are shown in Table 1. Marshall did not apply the SETI test to equations 3, 6, or 16, because he had already rejected them on the basis of his naive model tests. The SETI test obviously would have rejected them, however. In 1946 and 1947 they have by far the largest calculated disturbances in the model. Also, for each of these equations in 1946 and 1947 the disturbance is between 5 and 6 times as large as its *maximum* value in the sample period, and between 6 and 18 times as large as its estimated standard error.

Therefore, we conclude that by the SETI test equations 3, 6, and 16 are

[29] Milton Friedman too has suggested these naive models, though not under this name.

[30] Furthermore, the size of a structural equation's disturbance is not an invariant for this purpose because a structural equation can be normalized arbitrarily on any endogenous variable, but the size of a reduced-form equation's disturbance is a definite quantity because there is only one dependent variable on which to normalize a given reduced-form equation. Marshall comes close to realizing this when he comments that the verdict of a naive model test of a structural equation depends on which variable is selected from the equation as a basis for the test. He always chooses the one Klein has placed on the left side, and he does realize that this is an arbitrary choice. See Marshall (17).

Table 1

RESULTS OF MARSHALL'S SETI TEST OF KLEIN'S MODEL III

Eq.	Var. at left	Yr.	Calc. dist. u^*	ks_1^*	ks_4^*	ks_5^*	Verdict[a]
1	I	46	−5.6	2.0	3.4	8.0	
		47	−2.3		2.9	3.6	
2	H	46	−.7	2.6			
		47	−1.9				
3	ΔX	46	−61.0				[b]
		47	−37.9				
4	W_1	46	7.2	4.4	9.2	15.9	
		47	14.7		8.7	8.9	R
6	C	46	12.7				[b]
		47	14.0				
7	D_1	46	.3	.9	.9		
		47	−1.3		1.3		
8	v	46	1.5	3.6	4.5		
		47	4.6		4.4	8.1	
9	Δr	46	−.1	.1			
		47	−.0[c]				
10	D_2	46	.6	1.2			
		47	.5				
11	Δi	46	−.8	2.1			
		47	−.6				
16	M_1	46	1.9	5.3	6.4		
		47	−9.0		6.5	6.9	R
17	M_2	46	12.2				[b]
		47	12.0				

Source: Marshall (17). [a] R means reject; a blank space means accept.

[b] Marshall did not apply the SETI test to this equation because he rejected it on the basis of his naive model test.

[c] Less than .05 in absolute value, and negative.

rejected; equations 4 and 15 are on probation for having been rejected for either 1946 or 1947; and equations 1, 2, and 7-11 have a clear record so far.

Since neither Klein or Marshall made any explicit computations of the reduced form, results of naive model tests of the reduced form of Klein's model III are not presented here.

7 REVISIONS OF KLEIN'S EQUATIONS

This section presents several equations designed to replace those of Klein's which fared badly in Marshall's SETI test.

The SETI test, as indicated in Section 6, would have rejected three equations: (3) output adjustment, (6) demand for consumption, and (16) demand for idle cash balances. It cast doubt upon two others: (4) demand for labor and (15) demand for active cash balances. These five equations are the ones to be revised or changed here. If theoretically justified, it is permissible to change the number of variables and equations in the model, but the number of equations must not exceed the number of endogenous variables, and the two must be the same if the system is to be complete.

Consider first the demand for money equations (15 and 16). Their function is to determine two variables, $M_1{}^D$ and $M_2{}^D$ (active and idle money balances, respectively), which are purely symptomatic in Klein's model. Since they do not enter into any other equations, $M_1{}^D$ and $M_2{}^D$ cannot mathematically affect the other variables of the model but can only be affected by them. We are not interested in the quantity of money per se unless it has some effect. Therefore, we drop equations 15 and 16, together with the variables $M_1{}^D$ and $M_2{}^D$.[31] This cuts the number of equations by two but still leaves a complete model.

The demand for dwelling space, equation 8, is in the same position as the demand for money equations. It determines the nonfarm housing occupancy rate, v, whose current value does not appear elsewhere in the model. Therefore it can be dropped, along with variable v. We have now removed three equations and three variables jointly dependent at time t, without affecting the completeness of the model.

Consider the consumption function next. Klein's equation 6 underestimated consumption in 1946 and 1947 by some 13 and 14 billions of *1934* dollars, or about 15 and 16 per cent.[32] The real value of the stock of money at the beginning of each of these years was $110 and $105 billion, respectively, approximately twice the largest value attained during 1921-39. (The real value of the stock of money is here defined as the sum of currency outside banks plus demand deposits adjusted plus time deposits, but not

[31] I believe that a good economic theory will not say that the quantity of money is merely a symptom having no effect upon economic affairs. Accordingly, Klein's theory is amended below, at least with respect to the consumption function. Only lack of time prevented further changes involving the quantity of money in other parts of the model and dictated the dropping of 15 and 16 rather than their revision.

[32] See Marshall (17).

including government deposits, deflated by the 1934-base price index of output as a whole.) For the interwar years Klein (12) was unable to reject the hypothesis that consumption is not dependent upon real cash balances, but this was to be expected because real balances were almost constant during that period except for a smooth trend, so their effect, if any, could not be discovered. The postwar data suggest that real balances may have been important in the consumption function all along. The skewness of the distribution of ownership of real balances among consumers may also be important; we might expect to find that an increase in the holdings of richer people would stimulate consumption less than an equal aggregate increase in the holdings of poorer people (of course the same might be true of income). Time series are not available for this ownership distribution, however.

The proper definition of cash balances for this purpose is total consumer holdings of currency, demand deposits, time deposits, and probably also U. S. Savings Bonds (Series E) as long as they are guaranteed to be immediately redeemable in cash at no loss and yield negligible interest. Holdings by individuals and unincorporated businesses might be a good approximation, but suitable figures do not exist as far as I know, especially if Series E bonds are included. Therefore the definition used in the preceding paragraph seems best.

Lagged disposable income has often been mentioned as a candidate for membership in the consumption function. It is recommended by the fact that people do not adjust themselves immediately to changes in income. Lagged consumption has also been suggested, for a similar reason.[33] As lagged income and lagged consumption are highly correlated (through the consumption function) it is best not to use both.

Accordingly we experiment with fitting the following consumption functions:

$$(6.0)\qquad C = \delta_0 + \delta_1 Y + \delta_2 Y_{-1} + \delta_3 \left(\frac{M}{p}\right)_{-1} + \delta_4 t + u_6$$

$$(6.1)\qquad C = \delta_0' + \delta_1' Y + \delta_2' Y_{-1} + \delta_3' t + u_6'$$

$$(6.2)\qquad C = \delta_0'' + \delta_1'' Y + \delta_2'' \left(\frac{M}{p}\right)_{-1} + \delta_3'' t + u_6''$$

$$(6.3)\qquad C = \delta_0''' + \delta_1''' Y + \delta_2''' t + u_6'''$$

$$(6.4)\qquad C = \delta_0^{IV} + \delta_1^{IV} Y + \delta_2^{IV} C_{-1} + \delta_3^{IV} \left(\frac{M}{p}\right)_{-1} + \delta_4^{IV} t + u_6^{IV}$$

$$(6.5)\qquad C = \delta_0^{V} + \delta_1^{V} Y + \delta_2^{V} C_{-1} + \delta_3^{V} t + u_6^{V}$$

[33] The suggestion was made informally by Klein and by Franco Modigliani.

where M = currency outside banks + demand deposits adjusted + time deposits, at the end of the year; and other symbols are defined in Section 5. Observe that $(M/p)_{-1}$ is a predetermined variable since it is lagged (the same is of course true of Y_{-1} and C_{-1}). Thus we have not added any new current endogenous variables to the system by these modifications of the consumption function.

There remain two equations, (3) output adjustment (which is really a supply function, as mentioned before) and (4) demand for labor. They are closely related theoretically, because under the assumption of profit maximizing, the firm's demand-for-factor equations are deducible from the profit function and the production function; the supply function is then deducible from these demand-for-factor equations and the production function.[34] Equivalently, if the demand-for-factor equations and the supply equation are given, the production function is determined. Thus if we are concerned only with the logical completeness of the model, it does not matter whether it is the production function or the supply function that we include, provided the demand-for-factor equations are present.[35] We

[34] Suppose we are given competitive conditions, a production function

(1) $x = \phi(y_1, \cdots, y_n)$,

and a profit function

(2) $\pi = px - \sum_{1}^{n} q_i y_i$,

where x is output and p is its price, y_i is the input of a factor of production and q_i its price, $i = 1, \cdots, n$, and π is profit. Then the firm maximizes (2) with respect to the y_i, subject to the restraint (1), to get

(3) $p\dfrac{\partial \phi}{\partial y_i} - q_i = 0, \qquad i = 1, \cdots, n.$

If the set of simultaneous equations (3) is solved for the y_i, $i = 1, \cdots, n$, the result is the demand-for-factor equations

(4) $y_i = f_i\left(\dfrac{q_1}{p}, \cdots, \dfrac{q_n}{p}\right), \qquad i = 1, \cdots, n.$

The supply equation is obtained by substituting y_i from (4) into (1), $i = 1, \cdots, n$. Results are similar in the noncompetitive case, but elasticities of product demand and factor supply enter in then.

[35] Under the assumption of profit-maximizing, with a profit function such as (2) in the preceding note, and with a set of demand-for-factor equations for the firm that can be uniquely solved for the factor prices, the production function for the firm can be derived from given demand-for-factor equations, uniquely except for a boundary condition such as $\phi\ (0, \cdots, 0) = 0$, even with no knowledge of the supply function, as follows: By hypothesis it is possible to pass uniquely from (4) to (3) of the preceding note, which can then be divided through by p and integrated to obtain ϕ uniquely except for a constant term (subject to certain integrability conditions which in our case are satisfied), Q.E.D. This proof is due to Koopmans.

This system is not likely to be made overdetermined by including a production function (or alternatively a supply function), however, since an additional variable x is brought along at the same time.

choose here to use a production function, because Klein's output adjustment equation is so far off (overestimating output by 61 and 38 billions of 1934 dollars in 1946 and 1947, respectively) and because the production function is less likely to be affected by possible structural changes.

Variables must be chosen to represent capital input and labor input in the production function. Capital input can be measured by depreciation charges, which would be ideal if depreciation really reflected the services of capital accurately. But since depreciation is a very arbitrary thing, subject to various legal and accounting pressures, it is not a satisfactory measure of capital use. Another possible measure is the stock of producers' capital at the beginning of the year, defined as the sum over time of net investment. This is not free from the effects of the arbitrariness of depreciation charges but it is less sensitive to them because stock of capital is so large in relation to depreciation charges for any one year. It measures capital *existing,* not capital in *use,* which is unfortunate, but we shall try it anyway, perhaps together with some device for indicating the extent to which available capacity is being used.

Labor input, which might appear also in the demand for labor equation, should ideally be measured in man-hours.[36] But data difficulties deter us here; the BLS series for average weekly working hours before 1932 is for manufacturing and railroads only, and does not cover all industries even now. The concept of full time equivalent persons engaged in production, used by Simon Kuznets and the Department of Commerce, is the next best thing. However, it does not regard overtime work as an increase of labor input: it measures roughly the number of persons engaged full time or more (where full time for any person means simply the current customary work week in his job, whether it is 35 hours or 55), plus an appropriate fraction of the number of persons engaged part time (to convert them to full time equivalents). A time trend term will then approximately take care of the secular decrease in weekly working hours that has occurred.[37]

We might choose any one of several forms for the production function. The Cobb-Douglas function, linear in logarithms, is one possibility; a simple linear or quadratic function is another. Investment during the current year might be included on the theory that new capital, because of

[36] It is private labor input that concerns us here, by the way, not total, because only in the private sector is production assumed to be guided by the desire for profit.

[37] Cyclical fluctuations in weekly working hours will be an important source of error here unless their effect is largely explainable by cyclical changes in full time equivalent persons engaged plus a time trend, i.e., if data on weekly working hours (which we do not have), full time equivalent persons engaged, and time trend are not approximately linearly related.

improvements in the design of equipment, is more productive than old capital even after depreciation has been deducted.

In attempting to make the production function reflect the fact that output can be increased if existing capital is used more intensively, we might break our sample into two samples — one containing boom years in which capital was being used at approximately full capacity, and the other containing slack years in which it was not — and then fit two production functions, one to each. The sign of net investment could be used as a crude indicator for classifying the years: in boom years one would expect demand for capital services to exceed existing supply, thereby stimulating an increase in the stock of capital, so that net investment would be positive, and in slack years the opposite. This scheme is undesirable because it sets up a dichotomy where there should be a continuum, and because it reduces the already too small sample. An alternative, suggested during discussions with Jacob Marschak, is to make each parameter of the production function a linear function of net investment, thus:[38]

$$X = (\mu_0 + \mu_1 I) + (\mu_2 + \mu_3 I)N + (\mu_4 + \mu_5 I)K_{-1} + \mu_6 t + u_3 \tag{3.0}$$

where N = private labor input, in millions of full time equivalent man-years (endogenous), and other variables are defined in Section 5. We would expect μ_5 to be positive: a large positive net investment I can be presumed to indicate that capital is being used at a high percentage of capacity, and existing capital K_{-1} can be expected to contribute more to output than otherwise, so that its coefficient $(\mu_4 + \mu_5 I)$ should be high. We might expect μ_3 to be negative because the marginal product of labor is probably less in boom times than otherwise. Of course we expect μ_2, μ_4, and μ_6 to be positive (though μ_6 would be negative if the above mentioned secular drop in working hours were enough to overbalance the increase in per man-hour productivity). We have no presumptions about μ_0 and μ_1, except that μ_1 should probably be positive and not very important.

Another way of trying to solve the problem of unused capacity is to set up a production function in which output depends upon both labor input and existing capital in boom years, but only upon labor input in slack years. This again unfortunately requires a dichotomous classification of all years as either boom or slack. Mainly because of lack of time, estimates of this kind of production function are not presented in this paper; it would be interesting to return to this idea in the future.

[38] Such a production function is dependent upon the assumption that net investment occurs in response to near-capacity use of existing capital. If something happens so that this is no longer true, the production function changes. But nothing of this sort is likely to happen unless the profit maximizing assumption becomes invalid, in which case several other equations will go by the board too.

Besides 3.0, we try the following production functions:

$$(3.1)\qquad X = \mu_0' + \mu_1'I + \mu_2'N + \mu_3'K_{-1} + \mu_4't + u_3'$$

$$(3.2)\qquad X = \mu_0'' + \mu_1''N + \mu_2''K_{-1} + \mu_3''t + u_3''$$

$$(3.3)\qquad \log X = \mu_0''' + \mu_1'''\log N + \mu_2'''\log K_{-1} + \mu_3'''t + u_3'''$$

$$(3.4)\qquad X = \mu_0^{IV} + \mu_1^{IV}N + \mu_2^{IV}t + u_3^{IV}$$

$$(3.5)\qquad X = \mu_0^{V} + \mu_1^{V}N + \mu_2^{V}NK_{-1} + \mu_3^{V}K_{-1} + \mu_4^{V}t + u_3^{V}$$

$$(3.6)\qquad X = \mu_0^{VI} + \mu_1^{VI}N + \mu_2^{VI}N^2 + \mu_3^{VI}NK_{-1} + \mu_4^{VI}K_{-1}^2 + \mu_5^{VI}K_{-1} + \mu_6^{VI}t + u_3^{VI}$$

In 3.1 to 3.4 we would expect all parameters (except possibly the μ_0's) to be positive. In 3.5 and 3.6 we would expect μ_2^V, μ_4^V, μ_3^{VI}, and μ_6^{VI} to be positive, and μ_1^V, μ_3^V, μ_2^{VI}, and μ_4^{VI} to be negative (this can be seen more easily by examining the expressions for marginal productivity of labor and capital implicit in the two equations).

Equation 3.3 is a Cobb-Douglas function with a time trend to take approximate account of technological improvements. 3.2 is a linear approximation. 3.1 is like 3.2 except that it treats new and old capital differently. 3.4 is a linear approximation which attempts to account for the existence of unused productive capacity by (1) disregarding the quantity of *existing* capital and (2) assuming (more or less plausibly) that capital input (not measurable) is proportional to labor input, so that output can be expressed as a function of labor input alone. 3.5 and 3.6 are attempts at more accurate approximation than a linear function provides: they have marginal productivity functions that vary with inputs instead of being constant.

Observe that by replacing Klein's output adjustment equation with any of the production functions 3.0 to 3.6, we have added a new endogenous variable, N. Before we finish our revisions, we must therefore find a corresponding additional equation, if we are to end with a complete system.

Now that the wage-salary bill and labor input are both in the model, it is natural to include the wage rate too:

$$(18)\qquad w = \frac{W_1}{N}$$

where w = private wage-salary rate, in thousands of current dollars per full time equivalent man-year (endogenous), and W_1 and N are defined in Section 5 and in this section, respectively.

Adding 18 will not affect the completeness of the system. We have here one new equation and another new endogenous variable, w, so we still need to find an additional equation.

If the wage rate is to mean merely total labor earnings per unit of labor input, the definition is satisfactory in the simple form (18). However, if it is to mean the hourly wage, the thing over which workers and employers bargain, overtime payments, premiums for night-shift work, etc., must be allowed for; furthermore, labor input must be measured in man-hours. The advantage of using the hourly wage is that it enables us to introduce an equation describing the bargaining process and its dependence upon price movements, level of employment, and any other relevant variables. This wage adjustment equation could serve also as the additional one required by the introduction of the two new endogenous variables, w and N, with only the one equation 18. But existing data do not permit us to incorporate overtime payments and premiums for shift-work into the wage rate or, as we have seen, to define labor input in man-hours.

Accordingly we retain 18 as it stands. We assume that our w is closely representative of hourly wage,[39] and use a wage adjustment equation such as:

$$(5.0) \qquad w = \kappa_0 + \kappa_1 \Delta p + \kappa_2 (N_L - N) + \kappa_3 w_{-1} + \kappa_4 (N_L - N)_{-1} + \kappa_5 t + u_5$$

or

$$(5.1) \qquad w = \kappa_0' + \kappa_1' \Delta p + \kappa_2' (N_L - N) + \kappa_3' w_{-1} + \kappa_4' t + u_5'$$

where N_L = labor force, including work relief employees but excluding other government employees,[40] in full time equivalent man-years (exogenous),[41] and other variables are as defined above. These wage adjustment

[39] A study in the *Monthly Labor Review* for November 1942, pp. 1053-56, shows the estimated average number of overtime hours per worker per week in manufacturing in 1942 as a function of average total hours per worker per week. Using this study and the BLS series for average weekly hours in manufacturing, and assuming that the 1942 study is valid for all years and that all time over forty hours is paid at time and a half, one concludes that if overtime pay had been the sole cause of the difference between our w and the straight-time hourly wage, this difference would have been less than 2 per cent in all interwar years and less than 3 per cent in 1946 and 1947. Thus we are not risking more than about 3 per cent from this cause. Shift premiums probably do not contribute a larger error than this. And we are more comfortable if we remember that the manufacturing industries probably had more extensive shift premiums and more complete observance of the time and a half for overtime rule than did the economy as a whole.

[40] $N_L - N$ is meant to measure unemployment including relief workers, and N excludes government workers. Therefore if $N_L - N$ is to be a correct measure, N_L must exclude government nonrelief workers.

[41] Labor force is the only measure we have for labor supply, and it is expressible in man-years, but not in man-hours except by some trick assumption. Hence unemployment and employment, which add up to labor force, must also be in man-years instead of man-hours. Hence, there is another advantage in defining labor input N in man-years as we have done.

equations tell us that the wage level depends upon the past wage level (reflecting the downward rigidity of wages), upon price changes (reflecting wage increases following increases in the cost of living and in the prices received by employers), upon unemployment (reflecting the state of the labor market), and upon trend (reflecting the growth in productivity and/or in the strength of unions).

By adding a wage adjustment equation, we have completed our system again.

The demand for labor equation (4) is still to be considered. It was put on probation, not completely rejected. Therefore we try it again, but we also try two alternatives which express the demand for labor in real terms as a function of the real wage rate, of real output, and possibly of trend. The real wage rate enters as a result of the profit maximizing assumption. Output is relevant on the theory that if producers receive more orders they will demand more labor even if the real wage does not fall.[42] Trend may be necessary to reflect the long-term rise in per man-hour productivity. Our alternative equations are:

$$(4.0) \qquad W_1 = a_0 + a_1(pX - \mathcal{E}) + a_2(pX - \mathcal{E})_{-1} + a_3 t + u_4$$

$$(4.1) \qquad N = a_0' + a_1' \frac{w}{p} + a_2' X + a_3' t + u_4'$$

$$(4.2) \qquad N = a_0'' + a_1'' \frac{w}{p} + a_2'' X + u_4''$$

Klein's equation 4.0 is not as different from the others as it looks at first: if we divide 4.0 through by w we get

$$(19) \qquad \frac{W_1}{w} = N = \frac{a_0}{w} + \frac{a_1 X}{w/p} - a_1 \frac{\mathcal{E}}{w} + \cdots$$

which also depends on real wage w/p and on real output X, though there is only one parameter, a_1, to take care of both, and there are other terms involving w and $\mathcal{E}$ and lagged quantities meant to account for expectations.

Whether we finally choose 4.0 or 4.1 or 4.2, we still have a complete system of fourteen equations (including four definitional identities) in fourteen jointly dependent variables: $I, H, X, W_1, w, C, D_1, r, D_2, i, Y, p, K$, and N.[43] It should be understood that there are additional identities which define as additional variables the following nonlinear functions in

[42] The dependence of the demand for labor upon output cannot be found from the profit maximizing assumption and the usual production function, which may indicate a weakness in one of these two.

[43] Klein (13) at least hints at most of the changes made in this section, and even includes exploratory computations on some.

the system: $(pX-\mathcal{E})/q$, $pX-\mathcal{E}$, w/p, M/p, r/q_1, $1/r_{-1}$, and $(1/p)(W_2 + R_1 + R_2)$.

8 DESCRIPTION OF TESTS USED

Several tests are available for application to a model or structure obtained by the methods described in this paper. They may be divided into two groups according to the information required for their use. The first group comprises tests dependent only on observations and restrictions available for use in the estimation process; these are essentially tests of internal consistency. The second group comprises tests that use observations concerning events outside (usually subsequent to) the sample period; these are tests of success in extrapolation and prediction, and therefore are of higher authority. We describe here the tests applied in this paper.

a) *Tests of internal consistency*

First, there are certain qualitative procedures that perhaps should not be called tests at all: the estimates of the structural parameters can be examined to see whether they have the approximate magnitudes and particularly the algebraic signs to be expected on the basis of theoretical and other information about elasticities, marginal propensities, etc. The estimated sampling variance of each estimate can be examined to see how much confidence can be placed in its sign or in its approximate size. The calculated disturbances can be examined to see whether they are very large according to some intuitive standard of how large they are expected to be. This last procedure is of doubtful usefulness because it is not always possible to tell whether disturbances are due to the existence of several systematic factors that have been neglected, or to a real randomness in the phenomenon studied, especially if the disturbances appear to be random.

Second, for any equation of the model there is a test of all the restrictions used in the limited information estimation of that equation. The test is applied to the largest characteristic root λ_1 of the equation

(20) $$det[W(W^* - W)^{-1} - \lambda I] = 0$$

which is used in the estimation process. Here W is the covariance matrix of disturbances to the regressions *of* the H jointly dependent variables in the equation to be estimated *on* the predetermined variables assumed to be known to appear in the entire model; W^* is the covariance matrix of disturbances to the regressions *of* the same H jointly dependent variables in the equation to be estimated *on* the predetermined variables in the equation to be estimated; the roots $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_H$ are scalars; and I is the identity matrix. Anderson and Rubin (1) have shown that under the assumptions of the limited information method, the quantity $T \log (1 + 1/\lambda_1)$ has the χ^2 distribution asymptotically as the sample size T increases, with the number of degrees of freedom equal to the number of overidentifying restrictions, i.e., to the excess of K^{**}, the number of predetermined variables assumed to be known to enter the model but not the given equation, over $H-1$, where H is the number of jointly dependent variables in the given equation. $1 + 1/\lambda_1$ can never be less than 1, and if it is close to 1 in an overidentified model it means that the effect of excluding the excluded predetermined variables is only slightly detrimental to the variances, i.e., increases them only slightly, which is what we want. This χ^2 test of the largest root λ_1 is a sort of over-all test of the totality of restrictions and assumptions applied in estimating an equation; if in a particular equation, λ_1 takes a value that is very improbable under the hypothesis that all these assumptions are true, then for that equation we have only a very generalized alarm signal which cannot point to a specific remedy. (Of course, if we have a high degree of a priori confidence in some *specific* set of identifying restrictions, this test can be regarded as a test of the remaining, overidentifying, restrictions.) The test is of questionable usefulness for this paper because it is derived on the assumptions of the ordinary limited information method, and this paper uses the abbreviated variant of the limited information method.

Third, for any equation of the model there is a test of the assumption that the disturbances are serially uncorrelated, based on the distribution of the statistic δ^2/S^2. For a given equation and sample period, δ^2 is the mean square successive difference of the disturbances u, given by

$$(21) \qquad \delta^2 = \frac{1}{T-F-1} \sum_{t=2}^{T} (u_t - u_{t-1})^2$$

and S^2 is the variance of the disturbances u over the sample, given by

$$(22) \qquad S^2 = \frac{1}{T-F} \sum_{t=1}^{T} u_t^2$$

where T is the sample size and F is the number of parameters to be estimated in the equation. The question of the proper number of degrees of freedom is not solved, so we follow Marshall in using $T-F$ arbitrarily, as if we were dealing with least squares. The distribution of δ^2/S^2 has been tabulated by Hart and von Neumann (8).[44]

b) Tests of success in extrapolation and prediction

First, there is the SETI test, described in Section 6. As stated there, the SETI test tells whether each structural equation describes events in future

[44] Since we never *observe* the disturbances, being forced to calculate their values on the basis of estimates of the structural parameters, there may be some bias in using the tables given by Hart and von Neumann. In fact, Orcutt and Cochrane (18) have found in sampling experiments that there is a high probability of bias against finding serial correlation, especially when the number of parameters to be estimated is large.

periods as well as it does those in the past sample period. A similar test could be applied to the estimates of the equations of the reduced form; it might be called the reduced form tolerance interval test, or the RFTI test for short. The RFTI test is a test of *predictions,* but in case of poor prediction it cannot tell us which structural equations should be changed.[45] The SETI test examines each structural equation separately, and is therefore more useful in this respect.[46]

Second, there are the naive model tests of the predicting ability of the reduced form of the econometric model, mentioned in Section 6. We want the errors of prediction made by the reduced form, i.e., the calculated disturbances to the reduced form equation in the years for which predictions are made, to be at least as small in absolute value as the errors made by the noneconomic naive models.[47] If this condition is not met, we cannot have much confidence in the predicting ability of the econometric model. But observe that even if such a naive model does predict about as well as our econometric model, our model may still be preferable because it may be able to predict consequences of alternative policy measures and of other exogenous changes, while the naive model cannot.

Third, a comparison can be made to see whether the limited information method or the least squares method yields smaller calculated disturbances to the structural equations in the years for which predictions are made.

[45] 'Autonomy' of an equation is the name given to a concept that is useful here. It is not numerically defined, but corresponds to the degree to which the equation is invariant under possible or probable changes in structure. Structural equations are the most autonomous, since each depends on the structural parameters of only one equation, namely itself. Reduced form equations are the least autonomous. The advantage of autonomous equations is obvious for prediction under changes of structure. See Haavelmo (6).

[46] If a structural equation with limited information estimates of its parameters fails to pass the SETI test, we can be reasonably confident that the trouble (apart from sampling variation) lies with the form of *that* equation and not with the other equations of the model, because in estimating that equation no information from the rest of the model was used, except for observations on a list of predetermined variables. This statement could fail to be true only if the rest of the model contained a seriously wrong set of predetermined variables. But observe this caution: even if all the calculated disturbances fell inside their tolerance intervals, we still might not have a good structure; we might instead have a poor structure which, however, is not worse in the prediction period than it was in the sample period (this remark arose in discussions with Harry Markowitz).

[47] To test this, we can make point predictions (as opposed to tolerance interval predictions) with both methods for a number of years, and apply a simple *t*-test to the hypothesis that the means of the absolute values of their errors are the same, using as an alternative the hypothesis that the mean of the absolute values of the econometric model's errors is larger. As we are likely to have very small samples for this test as well as for the SETI test (Marshall would have had a sample of two, for instance) its results will not be conclusive.

9 PLAN OF COMPUTATIONS

Klein has estimated the equations of his model (Sec. 5) by the least squares and limited information methods; the estimates are given in Klein (13).[48] Certain of these equations have been rejected by Marshall's SETI test on the basis of Klein's limited information estimates and the data for 1946 and 1947; the results are given in Section 6 above. The rejected equations have been revised, replaced, or eliminated (see Sec. 7).

The estimates presented here are for the unrejected Klein equations, and for the new or revised equations of Section 7. They are based in each case on a sample consisting of the years used by Klein for his limited information estimates plus 1946 and 1947,[49] which were added in order to bring the estimates up to date and give the model a fairer chance to do a good job of describing 1948. The war years 1942-45 were omitted because some of the ordinary economic relationships were set aside in favor of direct government controls during that period. Some controls, e.g., rent controls, continued after 1945, however, and some period of readjustment may be required before the postwar economy finds its stride. After a few years, when the sample of postwar years has grown, it may be wise to drop 1946 as well as 1942-45.

All the unrejected and new equations are estimated by least squares, and the estimated standard errors of the disturbances and of the estimates are computed. Then *one* form is chosen from the theoretically acceptable alternative forms of each equation, e.g., one production function from equations 3.0 to 3.6, etc., and estimated by the limited information method.[50]

[48] The estimates appearing in Klein (11) have been revised because of the discovery of an error in the time series for X. The revised series is used in Klein (13) and in this paper.

[49] This means that my sample is 1922-41 plus 1946-47 for all equations except 10.0 and 11.0, for which it is 1921-41 plus 1946-47.

See Appendix C for a discussion of certain peculiarities in the time series obtained for 1946 and 1947.

[50] The choice is based partly on theoretical grounds (but not wholly, or else it could be made before any empirical work is done), and partly on the least squares estimates. There is a presumption that if an equation fits well by least squares, i.e., if its residuals and the estimated standard errors of the estimates of parameters are small, there is likely to be a relation among its variables that can be consistently estimated by the limited information method. This is particularly true if the variance of the disturbance to the equation is small; see Jean Bronfenbrenner (2). I realize that this procedure is not satisfactory to the uncompromising advocate of consistency in estimation. Ideally all the alternative forms of each equation should be estimated by the limited information method, but as this is an expensive process the least squares estimates are used as a kind of screening device. How misleading they can be is shown in the cases of equations 1.0 and 4.0, discussed below.

For each equation estimated by the limited information method, estimates are prepared for: the standard error of the disturbance; the covariances of the estimates of the parameters; the successive values of ks_i^* required for the SETI test, where $P = .95$ and $\gamma = .99$; the value of the ratio δ^2/S^2; and the quantities needed for the characteristic root test. The calculated disturbances for 1948 are computed for Klein's limited information estimates, for my least-squares estimates, and for my limited information estimates. The SETI test is applied to the last.

The parameters of the reduced form are estimated by the ordinary least-squares method and by the restricted least-squares method.[51] The naive model tests are applied to both sets of estimates, with the single year 1948 as a sample.

The results of all these computations appear in the next section.

10 RESULTS OF COMPUTATIONS

Table 2 shows the computational results that are applicable directly to structural equations (as opposed to equations of the reduced form): the estimates of parameters and variances, the calculated 1948 disturbances, and the quantities needed for the SETI test, the serial correlation test, and the characteristic root test.

Table 3 presents results pertaining to the equations of the reduced form of the revised model. For each of the endogenous variables,[52] it shows: (1) the observed 1948 value; (2) the change in the observed value from 1947 to 1948; (3) the average absolute value of the annual change in the observed value, over the 24 periods 1920-21 to 1940-41 and 1945-46 to 1947-48; (4) and (5) the two 1948 predictions made by the reduced form of the revised model, as estimated by the ordinary least squares method and by the restricted least squares method, respectively; (6) and (7) the 1948 predictions made by naive models I and II; (8) and (9) the errors of the two reduced form predictions, i.e., the observed values minus the predicted values; (10) and (11) the errors of the naive models; (12) to (15) a comparison of each reduced form error with each naive model error, to see in each case which is smaller in absolute value; (16) the percentage error of the least squares prediction, using the 1948 observed

[51] In this paper each reduced form equation includes only the predetermined variables that appear in the corresponding group of structural equations. See Appendix D for the grouping and Section 4 for remarks about the properties of the estimates.

[52] Endogenous in the sense of the y_i' in section 2. The equations of the reduced form are *linear* regressions on certain predetermined variables. Therefore, the predicted value of a nonlinear function such as w/p cannot be expected to be the same when obtained from the quotient of the predictions of w and p as when obtained directly from a regression. Predictions of such nonlinear functions are not presented here.

Table 2: RESULTS OF COMPUTATIONS ON STRUCTURAL EQUATIONS OF REVISED MODEL[a]

Eq.	(1)	Estimates of Parameters and Standard Errors[b] (2)	(3)	(4)	(5)	(6)	(7)	S (8)	δ^2/S^2 (9)	Obs. value 1948 (10)[c]	Calc. value 1948 (11)[c]	Calc. dist. 1948 (12)[c]	$ks_1^* = kS$ (13)[d]	1948 SETI test result (14)[e]	Pb of getting a smaller λ_1 (15)[f]
1.0 Investment	I	$\frac{pX-\mathcal{E}}{q}$	$\left(\frac{pX-\mathcal{E}}{q}\right)_{-1}$	K_{-1}	1										
KLS	1	.08 (.02)	.07 (.02)	−.12 (.02)	5.24			.42	1.81	1.89	4.82	−2.93			
KLI	1	.12 (.023)	.043 (.023)	−.098 (.023)	2.59 (2.7)			.47	1.59	1.89	5.39	−3.50			
CLS	1	.130 (.04)	−.041 (.03)	−.040 (.05)	.18			1.01		1.89	3.15	−1.26			
CLI	1	1.271 (2.0)	−.846 (1.4)	.52 (.32)	−78.38 (400)			6.62	1.50	1.89	10.26	−8.37	21.72	A	.45
2.0 Inventory	H	$X-\Delta H$	p	H_{-1}	1										
KLS	1	.13 (.02)	4.66 (1.2)	.48 (.08)	1.06			.57	2.17	34.34	36.86	−2.52			
KLI	1	.12 (.02)	4.59 (1.1)	.50 (.07)	1.17 (1.6)			.58	2.26	34.34	36.58	−2.24			
CLS	1	.110 (.02)	1.42 (1.3)	.467 (.11)	6.08			.82		34.34	33.16	1.18			
CLI	1	.082 (.02)	1.93 (1.4)	.539 (.12)	5.22 (1.9)			.87	1.64	34.34	33.15	1.19	2.86	A	.30
3.0[g] Production	X	I	IN	N	IK_{-1}	K_{-1}	t								
CLS	1	19.8 (2.0)	−.40 (.19)	3.966 (.84)	−.06 (.15)	−.146 (.24)	.29 (.31)	3.05		90.26	98.49	−8.23			
3.1 Production	X	I	N	K_{-1}	t	1									
CLS	1	.37 (1.38)	2.657 (.66)	−.303 (.17)	.62 (.29)	−19.75		3.33		90.26	98.54	−8.28			

3.2[h] Production	X	N	K_{-1}	t	1										
CLS	1	2.820	−.296	.55	−26.50			3.23		90.26	99.02	−8.76			
		(.25)	(.16)	(.14)											
*CLS**	1	2.812	−.426	.63	−12.18			2.99		90.26	96.74	−6.48			
		(.22)	(.16)	(.17)											
*CLS***	1	2.956	−.482	.70	−11.77			2.90		90.26	95.43	−5.17			
		(.19)	(.16)	(.14)											
3.3 Production	$\log X$	$\log N$	$\log K_{-1}$	t	1										
CLS	1	2.119	−.313	.003	−1.030			.006		1.956	2.025	−.070			
		(.17)	(.28)	(.001)											
3.4 Production	X	N	t	1											
CLS	1	3.074	.43	−68.13				3.42		90.26	99.16	−8.90			
		(.22)	(.13)												
CLI	1	2.917	.48	−61.93				3.47	1.52	90.26	98.08	−7.82	11.18	A	.08
		(.24)	(.13)	(9.4)											
3.5 Production	X	N	NK_{-1}	K_{-1}	t	1									
CLS	1	−.928	.034923	−1.706	.63	124.83		3.23		90.26	98.84	−8.57			
		(3.7)	(.035)	(1.4)	(.16)										
3.6[i] Production	X	N	N^2	NK_{-1}	K^2_{-1}	K_{-1}	t								
CLS	1	12.234	−.095663	−.015062	.042003	−8.708	.77	3.04		90.26	91.83	−1.57			
		(8.5)	(.050)	(.049)	(.031)	(7.0)	(.17)								
4.0 Labor	W_1	$pX - \mathcal{E}$	$(pX - \mathcal{E})_{-1}$	t	1										
KLS	1	.47	.12	.19	4.70			.96	1.27	114.65	107.31	7.34			
		(.03)	(.03)	(.04)											
KLI	1	.41	.17	.17	5.04			1.00	1.89	114.65	104.67	9.98			
		(.03)	(.03)	(.04)	(1.0)										
CLS	1	.480	.158	.20	**2.08**			1.47		114.65	112.63	2.02			
		(.04)	(.04)	(.05)											
CLI	1	−8.286	8.949	3.49	15.17			76.2	1.36	114.65	24.09	90.56	249.9	A	.10
		(127)	**(128)**	**(48)**	**(195,000)**										

Eq.	(1)	Estimates of Parameters and Standard Errors[b] (2)	(3)	(4)	(5)	(6)	(7)	S (8)	δ^2/S^2 (9)	Obs. value 1948 (10)[c]	Calc. value 1948 (11)[c]	Calc. dist. 1948 (12)[c]	$ks_1^* = kS$ (13)[d]	1948 SETI test result (14)[e]	*Pb* of getting a smaller λ_1 (15)[f]
4.1 Labor	N	w/p	X	t	1										
CLS	1	−8.46 (8.5)	.344 (.05)	−.06 (.06)	28.52			1.06		52.07	49.42	2.65			
4.2 Labor	N	w/p	X	1											
CLS	1	−13.81 (6.3)	.364 (.05)	32.09				1.06		52.07	49.93	2.14			
CLI	1	−23.08 (7.9)	.456 (.06)	35.20 (3.9)				1.19	1.09	52.07	51.30	.77	3.83	A	.04
5.0 Wage	w	Δp	$N_L - N$	w_{-1}	$(N_L - N)_{-1}$	t	1								
CLS	1	.54 (.11)	−.022 (.004)	.75 (.08)	.008 (.006)	.007 (.002)	.37	.019		2.20	2.06	.14			
5.1 Wage	w	Δp	$N_L - N$	w_{-1}	t	1									
CLS	1	.64 (.08)	−.018 (.002)	.67 (.03)	.01 (.002)	.48		.020		2.20	2.03	.17			
CLI	1	.81 (.13)	−.011 (.004)	.73 (.05)	.005 (.002)	.37 (.07)		.026	1.57	2.20	2.03	.17	.085	A[j]	.05
6.0 Consumption	C	Y	Y_{-1}	$(M/p)_{-1}$	t	1									
CLS	1	.558 (.08)	.053 (.10)	.270 (.07)	−.24 (.11)	6.70		1.60		82.84	75.76	7.08			
6.1 Consumption	C	Y	Y_{-1}	t	1										
CLS	1	.502 (.10)	.361 (.09)	.08 (.10)	4.67			2.24		82.84	74.46	8.39			

6.2 Consumption	C	Y	$(M/p)_{-1}$	t	1									
CLS	1	.583	.297	−.27	7.07		1.57		82.84	75.97	6.87			
		(.06)	(.04)	(.10)										
CLI	1	.543	.315	−.27	8.56		1.59	.91	82.84	75.96	6.88	5.22	R^k	.08
		(.06)	(.05)	(.10)	(2.9)									
6.3 Consumption	C	Y	t	1	Yt									
KLS	1	.77	.76	9.70	−.01		1.17	1.46	82.84	70.58	12.26			
		(.04)	(.06)		(.004)									
KLI	1	.73	.04	11.87	....		1.36	1.20	82.84	70.83	12.01			
		(.04)	(.06)	(2.5)										
CLS	1	.811	.21	7.31	....		2.88		82.84	75.66	7.18			
		(.08)	(.13)											
6.4 Consumption	C	Y	C_{-1}	$(M/p)_{-1}$	t	1								
CLS	1	.512	.214	.208	−.21	3.80	1.42		82.84	77.52	5.32			
		(.06)	(.09)	(.06)	(.10)									
6.5 Consumption	C	Y	C_{-1}	t	1									
CLS	1	.512	.462	.03	.08		1.39		82.84	79.21	3.63			
		(.07)	(.08)	(.09)										
7.0 Owned housing	D_1	r/q_1	$\Sigma_{-2}^{\circ} Y$	ΔF	1									
KLS	1	3.14	.02	.0039	−7.49		.21	1.80	1.86	4.93	−3.07			
		(.52)	(.003)	(.001)										
KLI	1	3.74	.024	.0043	−9.03		.21	2.26	1.86	5.07	−3.21			
		(.50)	(.003)	(.001)	(1.2)									
CLS	1	1.80	.0061	.0015	−2.65		.36		1.86	1.87	−.02			
		(.8)	(.004)	(.0005)										
CLI	1	1.79	.0072	.0014	−2.77		.36	.80	1.86	1.83	.03	1.17	A	.01
		(.9)	(.004)	(.0005)										

Eq.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	S (8)	δ^2/S^2 (9)	Obs. value 1948 (10)[c]	Calc. value 1948 (11)[c]	Calc. dist. 1948 (12)[c]	$ks_1^* = kS$ (13)[d]	1948 SETI test result (14)[e]	*Pb* of getting a smaller λ_1 (15)[f]
		Estimates of Parameters and Standard Errors[b]													
9.0 Rent	Δr	Y	v_{-1}	$1/r_{-1}$	1										
KLS	1	.0013	(.02)	.16	−2.13			.03	1.09	.066	.110	−.044			
		(.0007)	(.002)	(.04)											
KLI	1	.00071	.020	.17	−2.15			.03	1.04	.066	.051	.015			
		(.0007)	(.002)	(.04)	(.20)										
CLS	1	−.0003	.018	.16	−1.81			.033		.066	.059	.007			
		(.0008)	(.003)	(.06)											
CLI	1	−.0006	.018	.17	−1.85			.033	.96	.066	.055	.011	.109	A	.04
		(.0008)	(.003)	(.06)	(.27)										
10.0 Rental housing	D_2	i	r_{-1}	$(q_1)_{-1}$	$(q_1)_{-2}$	ΔF_{-1}	1								
KLS	1	−.25	2.93	.16	−.44	.0013	−1.99	.25	2.16	1.44	2.03	−.59			
		(.07)	(.39)	(.5)	(.5)	(.0006)									
KLI	1	−.18	2.81	.02	−.44	.0016	−2.14	.26	2.07	1.44	2.01	−.56			
		(.07)	(.34)	(.5)	(.5)	(.0005)	(.40)								
CLS	1	−.25	3.05	.04	−.53	.0011	−1.79	.25		1.44	1.59	−.14			
		(.06)	(.36)	(.5)	(.5)	(.0005)									
CLI	1	−.20	2.89	.0007	−.47	.0012	−1.92	.25	1.86	1.44	1.63	−.19	.85	A	.40
		(.07)	(.38)	(.5)	(.5)	(.0005)	(.43)								
11.0[i] Interest	Δi	ε_B	i_{-1}	t	1										
K	1	−.17	−.37	−.005	2.00			.47	1.77	.22	.70	−.48			
		(.10)	(.17)	(.035)											
C	1	−.092	−.34	−.03	**1.69**			.46	1.73	.22	.15	.07	1.49	A	l
		(.07)	(.16)	(.03)											

Notes to Table 2

[a] The equations are numbered here just as they are in the text except that equations 1, 2, 4, 7, 9, 10, and 11 in the text appear here as 1.0, 2.0, · · · , 11.0.

The variables in each equation are listed in the same row with the equation number, jointly dependent variables first and predetermined variables next. Each equation has a short title.

The units in which each variable is measured are given in Section 5 except that $\mathcal{E}_R$ is here converted to billions of current dollars, so that all quantities whose dimensions are current or 1934 dollars are measured in billions.

KLS and *KLI* refer to Klein's estimates by the least squares and limited information methods, respectively, based on a sample period ending in 1941, and found in Klein (13). (The sample was 1921-41 for all *KLS* equations except 1.0 and 9.0, and for *KLI* equations 10.0 and 11.0; the sample was 1922-41 for all other *KLI* and *KLS* equations.)

CLS and *CLI* refer to my estimates by the least squares and limited information methods, respectively, based on the *KLI* sample plus 1946 and 1947. (Thus the sample was 1921-41 and 1946-47 for *CLS* and *CLI* equations 10.0 and 11.0, and 1922-41 and 1946-47 for all other *CLS* and *CLI* equations.)

In the interest of not wasting effort in accurate computation of small quantities which will be added to larger and less accurate ones, relatively few significant figures are given for estimates of parameters attached to variables having small numerical values.

[b] The numbers in parentheses in columns 1-7 are estimates of the standard errors of the estimates of the parameters. The numbers not in parentheses are the estimates of the parameters. They are arranged in such a way that any equation may be read off directly in the form in which it is given in the text. For example, the *CLI* estimate of the consumption equation 6.2 is seen to be

$$C = .543Y + .315\,(M/p)_{-1} - .27t + 8.56.$$

[c] Column 10 gives the observed 1948 value of the variable appearing on the left side of each equation, i.e., the variable in column 1 of the table. Column 11 gives the value of the linear combination on the right side of each equation. Column 12 is column 10 minus column 11, the calculated disturbance. If this is positive, the equation has underestimated the variable on its left side. Column 10 minus 11 may not equal column 12 exactly because of rounding.

[d] The values of k for $\gamma = 0.99$ and $P = 0.95$, from Eisenhart, Hastay, and Wallis (3), p. 102, are as follows:

d.f.	*k*	*d.f.*	*k*	*d.f.*	*k*
15	3.507	18	3.279	21	3.121
16	3.421	19	3.221	22	3.078
17	3.345	20	3.168	23	3.040

[e] A = accept; R = reject (*CLI* equations only).

[f] Column 15 gives the approximate probability of obtaining a value of λ_1 smaller than was in fact obtained. A low probability indicates that our confidence in the *a priori* restrictions imposed must be low. See Table 4 for more details.

[g] The constant term in the *CLS* estimate of equation 3.0 is –84.76.

[h] *CLS** and *CLS*** under equation 3.2 refer to some special exploratory computations based on different series for K_{-1}. They are discussed in the text below.

[i] The constant term in the *CLS* estimate of equation 3.6 is 233.45.

[j] The value of ks^* for equation 5.1 is .181, which is larger than .17, the calculated disturbance. Hence the verdict is acceptance.

[k] The value of ks^* for equation 6.2 is 6.39, which is smaller than 6.88, the calculated disturbance. Hence the verdict is rejection.

[l] The *LS* and *LI* estimates of equation 11.0 are identical since it has only one dependent variable. No value of λ_1 is available for this equation.

---

value as a base; (17) the least-squares predicted change from 1947 to 1948; (18) a notation as to whether this predicted change was in the right direction.

11 DISCUSSION OF RESULTS OF COMPUTATIONS[53]

We shall look first at the results of (a) the SETI test and (b) the naive model tests. Also we shall (c) compare the 1948 calculated disturbances obtained from different estimates of each structural equation. Then we shall go back and look at the results of the tests of internal consistency described in the first part of Section 8: (d) the serial correlation test; (e) the characteristic root test; and (f) the qualitative examination of the estimates, in particular those for equations 1.0 and 4.0, where anomalous results appear.

a) There are ten stochastic equations in our revised econometric model, namely those estimated by the limited information method with a sample including 1946-47: the CLI equations 1.0, 2.0, 3.4, 4.0 or 4.2, 5.1, 6.2, 7.0, 9.0, 10.0, and 11.0. All are accepted for 1948 by the SETI test with $P = 0.95$ and $\gamma = 0.99$, except for the consumption function 6.2. If P and γ are both relaxed to 0.95, only one additional equation, the wage adjustment equation 5.1, is rejected by the SETI test. Even if P and γ are both relaxed to 0.75, all equations except 3.4, 5.1, and 6.2 are accepted by the SETI test with room to spare.[54] This means that nearly every equation fits 1948 as well as could be expected on the basis of its performance during the sample period.

b) For 1948, each of the two naive models predicts 7 out of 13 endogenous variables better, i.e., has smaller errors, than do the equations of the reduced form as estimated by the ordinary least squares method. Naive model *I* predicts better in 15 cases out of 21 than the reduced form as estimated by the restricted least-squares method, and naive model II pre-

[53] The results of all the computations in this paper of course depend upon the time series used for the variables for 1946 and 1947. See Appendix C for a discussion of certain peculiarities in those time series.

[54] To verify this, compare Table 2, columns 8 and 12, and Eisenhart, Hastay, and Wallis (3), p. 102.

Table 3

RESULTS OF NAIVE MODEL TESTS

Variable[a]	Eq.	Obs. Value 1948	Change in Obs. Value 1947-48	Average Absolute Change in Obs. Value[b]	Predictions 1948				Errors 1948[d]				Naive Model Test Verdicts[e]				Predictions by *LS* Estimates of Reduced Form		
					Reduced Form[c]		Naive Models		Reduced Form[c]		Naive Models		*LS*		*RLS*[c]		Prediction Error as % of (1)	Predicted Change 1947-48	Prediction of Direction of Change
					LS	*RLS*	*I*	*II*	*LS*	*RLS*	*I*	*II*	*I*	*II*	*I*	*II*			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Investment I	1.0	1.89	−.38	1.02	1.00	.98	2.27	3.22	.89	.91	−.38	−1.33	N	RF	N	RF	47	−1.27	right
Inventories H	2.0	34.3	3.04	1.25	27.4	34.4	31.3	31.6	6.9	−.1	3.0	2.7	N	N	RF	RF	20	−3.9	wrong
Price level p	2.0 5.1	2.03	.15	.07	2.17	2.24 2.17	1.88	2.19	−.14	−.21 −.14	.15	−.16	RF	RF	N RF	N RF	−7	.29	right
Private output X	3.4 4.2	90.3	2.96	5.03	72.5	96.7 79.5	87.3	79.3	17.8	−6.4 10.8	3.0	11.0	N	N	N N	RF RF	20	−14.8	wrong
Private employment N	3.4 4.2 5.1	52.1	1.05	1.81	48.7	51.6 49.4 45.9	51.0	53.5	3.4	.5 2.7 6.2	1.1	−1.4	N	N	RF N N	RF N N	7	−2.3	wrong
Private wage bill W_1	4.0	114.7	9.93	4.93	107.6	110.1	104.7	117.9	7.1	4.6	10.0	−3.2	RF	N	RF	N	6	2.9	right
Private wage rate w	5.1	2.20	.15	.07	2.10	2.08	2.05	2.22	.10	.12	.15	−.02	RF	N	RF	N	5	.05	right
Consumption C	6.2	82.8	1.13	2.86	80.2	67.8	81.7	76.9	2.6	15.0	1.1	5.9	N	RF	N	N	3	−1.5	wrong
Disposable income Y	6.2 7.0 9.0	79.9	1.74	4.15	70.9	65.0 78.4 70.9	78.2	66.4	9.0	14.9 1.5 9.0	1.7	13.5	N	RF	N RF N	N RF RF	11	−7.3	wrong
Owned housing D_1	7.0	1.86	.23	.23	2.04	1.62	1.63	1.92	−.18	.24	.23	−.06	RF	N	N	N	−10	.41	right
Rent level r	7.0 9.0	1.24	.07	.05	1.25	.81 .84	1.18	1.21	−.01	.43 .40	.06	.03	RF	RF	N N	N N	−1	.07	right
Rental housing D_2	10.0	1.44	.28	.22	1.49	1.73	1.16	1.38	−.05	−.29	.28	.06	RF	RF	N	N	−3	.33	right
Interest rate i	10.0 11.0	3.08	.22	.38	2.73	2.59 2.73	2.86	2.98	.35	.49 .35	.22	.10	N	N	N N	N N	11	−.13	wrong
Fraction of cases in which naive model error is less than reduced form error													7/13	7/13	15/21	13/21			

[a] Only the endogenous variables in the sense of the y_i' in Section 1 appear in this table. Each reduced form prediction is obtained from a reduced form equation containing all the predetermined variables in the group to which the predicted variable belongs. See Appendix D for the grouping.

[b] Average of the 24 periods 1920-21 to 1940-41 and 1945-46 to 1947-48.

[c] *LS* means least-squares estimates; *RLS* means restricted least-squares estimates.

[d] Error equals observed value minus predicted value. A positive error means underestimation.

[e] *RF* means that the reduced form's error is smaller than the naive model's error; *N* means the reverse.

dicts better in 13 cases out of 21 than the reduced form as estimated by the restricted least-squares method.[55]

These results do not permit us to say that there is any significant difference between the predicting abilities of the ordinary least squares estimates of the equations of the reduced form on the one hand and the naive models on the other. They suggest that, at least in the absence of structural change, predictions by the restricted least-squares estimates of the reduced form are inferior, both to predictions by the ordinary least-squares estimates of the reduced form and to those made by the naive models. The econometric model used here has failed, at least in our sample consisting of the one year 1948, to be a better predicting device than the incomparably cheaper naive models, even though almost every structural equation performs as well, i.e., has just as small an error, in extrapolation to 1948 as it does in the sample period.

It might be noted that the variables that are predicted better for 1948 by naive model *I* than by the reduced form (as estimated by either of the two ways) are almost exactly the same as those for which the change from 1947 to 1948 was less than the average (absolute value) of the annual changes over the sample period (see Table 3, col. 2, 3, 12, and 14). In other words, roughly speaking, naive model I predicted better the variables that changed *less* than usual, and the econometric model through its reduced form predicted better the variables that changed *more* than usual. This is not surprising, because naive model *I* assumes no change, and so of course will do well when there are only small changes, and poorly when there are large changes. On the other hand, the variables that are predicted better for 1948 by naive model *II* than by the reduced form include some for which the 1947-48 change was greater than average and some for which it was less (see Table 3, col. 2, 3, 13, and 15). But the variables whose *predicted* 1947-48 changes (based on the reduced form) were greater than average are not uniformly better predicted by the reduced form than by either naive model (see Table 3, col. 2, 12, 13, and 17). We conclude that it is not possible to tell in advance which variables are likely to be predicted better by the reduced form and which by a naive model.

However, an econometric model may be preferable, even though a naive model predicts equally well, because an econometric model may be able to predict the effects of alternative policy measures or other exogenous changes (including changes in structure if they are known about beforehand), while the naive model can only say that there will be no effect. Unfortunately we do not know how to tell rigorously in advance whether

[55] Incidentally, neither naive model is shown to be superior to the other; naive model II predicts better than naive model I in 7 out of 13 cases.

this will be true in a particular case, but it appears likely to be true when large or irregular changes occur in the exogenous variables, because it is then that the naive models are at their greatest disadvantage.

c) Table 2 shows that for every structural equation whose *KLI* and *CLS* estimates were both computed, the *CLS* estimates yield the smaller calculated disturbance. This suggests that for small samples a short extrapolation of least-squares estimates (i.e., from 1947 to 1948) may be more reliable than a prolonged extrapolation of limited information estimates (i.e., from 1941 to 1948). Table 2 shows also that for the eleven equations whose *CLI* and *CLS* estimates were both computed, the *CLS* estimates yield appreciably smaller calculated disturbances in four cases, the *CLI* yield smaller ones in two cases, and there is approximately a tie in five cases.[56] This suggests that short extrapolations based on least-squares estimates may be more reliable for samples as small as 22 than those based on limited information estimates.

We have here two comparisons of ordinary least-squares estimates with others known to be asymptotically superior (three if we recall that the ordinary least-squares estimates of the reduced form equations yield better predictions for 1948 than do the restricted least-squares estimates). In these comparisons the results suggest that in our problem the least-squares estimates lead to smaller errors in extrapolation. Now this is not surprising if there is no change in the underlying mechanism generating the observations, i.e., no change in structure. The argument is as follows: The least-squares method yields an estimate of the expected value of the conditional probability distribution of one variable, the one chosen to be "dependent", given the others. This distribution remains fixed as long as there is no change in structure. Therefore the least-squares estimates, which by construction produce the smallest possible calculated root-mean-square residual over the sample period, will continue to produce small residuals in extrapolation to subsequent periods as long as there is no change in structure.[57] But if the structure changes after the sample period and before the prediction period, the conditional probability distribution of the chosen dependent variable, given the others, will change in a complicated way, depending on the old and new structures. Then the least-squares estimates will no longer yield small errors in extrapolation, because they are estimates of the expected value of a distribution that is no longer relevant.

[56] The cases are, respectively: 1.0, 4.0, 9.0, 10.0; 3.4, 4.2; 2.0, 5.1, 6.2, 7.0, 11.0. Equation 11.0 *must* produce a tie because the *CLS* and *CLI* estimates are identical.

[57] The size of the error in extrapolation by any method will increase with the length of the extrapolation. For the case of least squares this is described by the Hotelling (9) formula for the standard error of forecast.

This is why it is desirable to estimate structural relations as well as simple regressions.

d) The limited information estimates presented here, as indicated in Section 1, are computed on the assumption that disturbances to the structural equations are not serially correlated.[58] For a sample of 22, if there is no serial correlation, the probability is 0.95 that the value of δ^2/S^2 will lie between approximately 1.26 and 2.93.[59] For the *CLI* estimates of equations 4.2, 6.2, 7.0, and 9.0, δ^2/S^2 is less than 1.26, indicating positive serial correlation of their disturbances.[60] Klein's limited information estimates give evidence of positive serial correlation of the disturbances to equations 6.3 and 9.0. There is no obvious relation between the performance of an equation in the SETI test and the serial correlation of its disturbances; no attempt has been made here to assess the error incurred by assuming zero serial correlation of disturbances.

Table 4

CHARACTERISTIC ROOT TEST RESULTS

small λ_1

Eq.	K^{**}[a] (1)	$H-1$[a] (2)	λ_1 (3)	$1+\frac{1}{\lambda_1}$	T (4)	$T \log_e (1+\frac{1}{\lambda_1}) = \text{col}(5) \times 2.3026$	$T \log_{10} (1 + 1/\lambda_1)$ (5)	$K^{**} - H + 1 = d.f.$[a] (6)	$\log_{10}$ Pb of getting a smaller λ_1 (7)	$\log_e$ Pb getting smaller
1.0	6	1	1.662	1.602	22	10.36	4.50	5	.45	.07
2.0	7	2	1.160	1.862	22	13.68	5.94	5	.30	.02
3.4	7	1	.450	3.222	22	25.74	11.18	6	.08	< .00
4.0	6	1	.605	2.653	22	21.46	9.32	5	.10	< .00
4.2	8	2	.345	3.899	22	29.93	13.00	6	.04	< .00
5.1	6	2	.587	2.704	22	21.87	9.50	4	.05	< .00
6.2	4	1	1.012	1.958	22	15.13	6.57	3	.08	.00
7.0	5	2	.109	10.174	22	50.98	22.14	3	.01	< .00
9.0	4	1	.686	2.458	22	19.80	8.60	3	.04	< .00
10.0	2	1	12.96	1.077	23	1.63	.71	1	.40	.20

[a] K^{**} is the number of predetermined variables that are assumed to be known to be in the model but not in the equation to be estimated; H is the number of jointly dependent variables in the equation to be estimated; and $K^{**} - H + 1$ is the number of overidentifying restrictions, i.e., the number of degrees of freedom of $T \log (1 + 1/\lambda_1)$; see Section 8.

e) Table 4, an expanded version of column 15 in Table 2, gives the results of the characteristic root test as applied to each equation of the revised model. At the 95 per cent significance level four equations are
all but 2

[58] Chernoff and Rubin have developed a consistent method of estimation, as yet unpublished, that does not require this assumption, but no computations have as yet been made with it.

[59] δ^2/S^2 is defined in Section 8, and its distribution is tabulated in Hart and von Neumann (8), p. 213. See also note 44 above.

[60] At the 90 per cent significance level the interval containing δ^2/S^2 is smaller, but no additional equations show serial correlation.

Geo. G. Judge
Asst. Prof Ag. Ec.
U. of Conn
Storrs, Conn.

And. Rev p. 56.

rejected by the characteristic root test: ~~4.2, 5.1, 7.0, and 9.0~~ 1.0 and 10.0. Furthermore, at the 90 per cent significance level ~~three~~ all but one of the other equations are rejected, ~~as well: 3.4, 4.0, and 6.2.~~ 10.0. Again, there is no obvious relation between the performance of an equation on this test and its performance on the SETI test, but this test rejects all equations rejected by the test of δ^2/S^2.

f) In the following discussion of the estimates presented in Table 2 we use abbreviated designations, such as "*KLS* 1.0" for "the Klein least squares estimates of equation 1.0."

In the demand for investment equation, *CLS* 1.0, the estimate of the coefficient of $\left(\frac{pX-\mathcal{E}}{q}\right)_{-1}$, which is closely related to lagged profits, has become negative, though not significantly. This may be due to sampling variation, or it may mean that entrepreneurs invest partly in response to *increases* in profits rather than only in response to high present and past profits. Thus, by means of the identity $\Delta x = x - x_{-1}$, *CLS* (1.0) can be equivalently written as

$$I = .089\left(\frac{pX-\mathcal{E}}{q}\right) + .041\Delta\left(\frac{pX-\mathcal{E}}{q}\right) - .040K_{-1} + .18. \tag{1.0'}$$

In *CLS* 3.0, 3.1, 3.2, and 3.3, i.e., in all the CLS production functions containing K_{-1} but no cross-product term in NK_{-1}, the marginal product of capital emerges as negative, though not significantly. At first this seemed to be due to the *fall* in K_{-1} from 110.1 in 1941 to 98.8 in 1946, coupled with the tremendous rise in X; it did not seem reasonable that the stock of productive capital in private hands had decreased 10 per cent during the war. But *CLS** 3.2 and *CLS*** 3.2, each based on an upward-revised postwar series for K_{-1},[61] yield even more strongly negative estimates of the marginal product of capital than *CLS* 3.2.[62] An examination of the

[61] The values of K_{-1} used in *CLS, CLS** and *CLS*** 3.2 are, respectively,

	K_{-1}	$K_{-1}{}^{*}$	$K_{-1}{}^{**}$
1946	98.787	102.600	111.400
1947	101.098	108.168	116.968

$K_{-1}{}^{*}$ is like K_{-1} except that additions are made to correct understatements during the war years due to the amortization of war plants in five years or less, allowed under the wartime revenue acts; the transfer of surplus producer goods from government to private hands; and (less defensible) the fact that the joint SEC-Department of Commerce series for plant and equipment expenditures, used in defining K, is smaller beginning in 1941 than the Department of Commerce series that appears in the national income accounts (the two series are almost identical before 1941). $K_{-1}{}^{**}$ is like $K_{-1}{}^{*}$ except that it assumes that the 1946 value is 111.4, as assumed in Klein (11), p. 135.

[62] *CLS** 3.2 and *CLS*** 3.2 have smaller calculated disturbances in 1948 than *CLS* 3.2 or *CLS* 3.4 or *CLI* 3.4, despite their negative marginal products of capital, because all the production functions have negative disturbances which are made smaller in magnitude by the presence of a larger negative term in stock of capital.

time series for X and K_{-1} shows that for about half of the sample period the two variables move in opposite directions; consequently, given the time series the result is not unreasonable. The conclusion to be drawn is either that the K_{-1} series does not measure stock of capital, as it is meant to, or that the stock of capital sometimes does not limit output. A less aggregative theory might be helpful in solving the problem. Equation 3.4 is the immediate solution chosen here.[63]

In *CLS* 3.0 the estimated coefficient of IK_{-1} is nearly zero when it is expected to be positive. But this is not a new cause for alarm, given the fact that the coefficient of K_{-1} is negative.

CLS 3.0 has one additional independent variable besides those in *CLS* 3.2, yet its disturbance has a larger estimated standard error, S. The same is true of *CLS* 6.0 and *CLS* 6.2, respectively, and of *CLS* 6.4 and *CLS* 6.5, respectively. This seems odd because when a new independent variable is added to a regression, it cannot increase the sum of squares of residuals. The answer lies in the fact that the reduction in the number of degrees of freedom caused by the introduction of the new variable more than uses up the reduction in the sum of squares brought about by the same cause. In such a case the additional variable is not worth its extra cost in degrees of freedom, except in larger samples where the cost is negligible.

The time trend term in the demand for labor equation *CLS* 4.1 has a very small coefficient not significantly different from zero, and therefore might reasonably be omitted. But the coefficient of w/p in *CLS* 4.1 and *CLS* 4.2 is very sensitive to the presence or absence of the trend term. As it too has a coefficient not significantly different from zero, however, its sensitivity might be attributed to sampling variation. It is apparent that in both 4.1 and 4.2 the chief relationship being estimated is that between X and N, namely the backbone of the production function. Indeed 3.4 is almost identical with either 4.1 or 4.2, numerical estimates and all — the term in w/p contributes relatively little to 4.1 or 4.2. It may be noted that 4.1 is not identified if 3.4 is in the model at the same time. Since 3.4 is the most satisfactory of our production functions, except for 3.5 and 3.6, which were tried later, we want to keep it, and so we replace 4.1 by 4.0 or 4.2.[64] Since all other equations meet the necessary condition (order con-

[63] The nonlinear equations 3.5 and 3.6, theoretically preferable to 3.2 and 3.4 because of having non-constant marginal productivities, were not estimated by the limited information method because of lack of time. Their least-squares estimates, particularly for 3.6, yield smaller calculated residuals for 1948 than any of the other production equations, however.

[64] This decision to drop 4.1 from the model is open to criticism because it is made in order to satisfy the necessary conditions for the identification of all equations, and not on theoretical or empirical grounds.

dition) for identifiability with room to spare, the probability is high that they meet the necessary and sufficient (rank) condition as well (see Sec. 1).

The consumption functions *CLS* 6.0, *CLS* 6.2, *CLI* 6.2, and *CLS* 6.4 show significantly positive coefficients for real cash balances $(M/p)_{-1}$. The addition of this term alone is enough to reduce the calculated disturbance in 1948 almost half — see *CLI* 6.2 — as compared with that of equation 6.3 which does not contain $(M/p)_{-1}$. But apparently the introduction of $(M/p)_{-1}$ is not sufficient to correct the consumption function, for *CLI* 6.2 is rejected by the SETI test. Another indication that $(M/p)_{-1}$ is not sufficient is that consumption has not fallen relative to disposable income since 1944, but $(M/p)_{-1}$ has been falling since 1946. Thus a term in $(M/p)_{-1}$ can explain the high postwar *average* level of consumption relative to income as compared with prewar, but it cannot explain the fact that consumption has remained high in 1947 and 1948, even exceeding disposable income, while $(M/p)_{-1}$ has been declining. It is evident that some other variable in addition to or in place of $(M/p)_{-1}$ is needed.

As can be seen from an examination of *CLS* 6.4 and *CLS* 6.5, lagged consumption expenditure C_{-1} appears to help matters, and more so when used *instead of* $(M/p)_{-1}$ than when used *in addition to* it.[65] *CLS* 6.5 has a smaller estimated standard error S and a smaller calculated 1948 disturbance than any of the other consumption equations estimated from the sample that includes 1946 and 1947.[66]

The estimated coefficients of disposable income Y in the rent adjustment equation *CLS* 9.0 and *CLI* 9.0 are negative. The explanation may lie in sampling variation, since the standard errors are of the same order of magnitude as the estimates. However, the controlled rise in postwar rents and the fall of Y from 1946 to 1947 may be responsible (see App. C).

The estimated coefficients of twice lagged construction costs $(q_1)_{-2}$ are negative in all four estimates of the demand for the construction of rental housing, equation 10.0. As their standard errors are about as large as the estimates, this need not be taken seriously. However, some response to expected costs, based on the past behavior of costs, may be indicated.

The *CLI* estimates of equations 1.0 and 4.0, as we have seen, are far out of line with our expectations. Unlike the *CLI* estimates of the other equations, they do not remotely resemble the *CLS, KLS,* and *KLI* estimates. Their calculated disturbances are often of the same order of magnitude as

[65] This is because the addition of $(M/p)_{-1}$ to 6.5 costs more in degrees of freedom than it is worth in reducing the sum of squares of residuals, as discussed in the third paragraph above.

[66] Limited information estimates were not computed for equations 6.4 and 6.5 because of lack of time; these equations were not considered until all the other computations were finished and the inadequacy of 6.2 became obvious.

the variables they contain, but they do not show clear signs of serial correlation.

One obvious possibility must be rejected immediately, namely that for each of these two equations the 1946 and 1947 observations may be nowhere near the line fitted to the 1922-41 sample, so that the line is radically changed by the addition of the 1946 and 1947 points. If this were true, the least squares estimates would be radically changed and the estimated standard error of disturbances greatly increased; but neither happens. Equations 1.0 and 4.0 are clearly cases where there is an approximately linear *empirical* relation among the variables (as evidenced by the least squares fits) but where the limited information method yields a straight line very different in slopes and intercepts from this empirical relation.

Sampling variation cannot be excluded as a possible explanation, especially since the estimated standard errors are so large that the *CLI* estimates do not differ *significantly,* i.e., by more than two or three times their respective standard errors, from the other estimates. Furthermore, nothing in the derivation of the limited information method requires it to yield small residuals and estimates close to the least-squares estimates, even though it has usually done so in the past.

There are two differences between the two procedures used in obtaining the *KLI* and the *CLI* estimates of equations 1.0 and 4.0. One is, obviously, that 1946 and 1947 are in the *CLI* sample but not in the *KLI* sample. The other is that the list of predetermined variables z^{**} (explained in App. D) for the *CLI* estimates differs from the list for the *KLI* estimates in that the variables X_{-1} and H_{-2} are omitted and the variables w_{-1} and $(N_L-N)_{-1}$ added. In other words, certain of the reduced form equations in the *CLI* case are regressions on a set of predetermined variables which differs from the corresponding set in the *KLI* case by containing w_{-1} and $(N_L-N)_{-1}$ instead of X_{-1} and H_{-2}, and accordingly the estimates of the parameters of equations 1.0 and 4.0 depend upon observations of a slightly different set of predetermined variables.

To separate the effects of these two changes, equation 4.0 was estimated four times: (1) (*KLI*) with the *Klein* z^{**}'s and *without* 1946-47; (2) with *my* z^{**}'s and *without* 1946-47; (3) with the *Klein* z^{**}'s and *with* 1946-47; (4) (*CLI*) with *my* z^{**}'s and *with* 1946-47. The results indicate that the anomalous *CLI* estimates of equation 4.0 are not due to the change in the list of predetermined variables z^{**}, but are somehow due instead to the addition of 1946 and 1947 to the sample.[67]

[67] (1) (*KLI*) $W_1 = .41\ (pX-\mathcal{E}) + .17\ (pX-\mathcal{E})_{-1} + .17t + 5.04$
(2) $W_1 = .413\ (pX-\mathcal{E}) + .175\ (pX-\mathcal{E})_{-1} + .17t + 5.05$
(3) $W_1 = 8.17\ (pX-\mathcal{E}) - 7.56\ (pX-\mathcal{E})_{-1} - 2.68t - 9.42$

12 SUMMARY AND CONCLUSION

The revised version of Klein's model, consisting of equations 1.0, 2.0, 3.4, 4.0 or 4.2, 5.1, 6.2, 7.0, 9.0, 10.0, 11.0, and the identities 12, 13, 14, and 18, has been subjected to several tests.[68] Table 5 summarizes the results of tests pertaining to the structural equations. Table 3 presents the results of the naive model tests, which pertain to the equations of the reduced form.

Table 5

SUMMARY OF RESULTS OF STRUCTURAL EQUATION TESTS

	Equation	SETI Test[a] $P=.95$ $\gamma=.99$	$P=.75$ $\gamma=.75$	Smaller Calc. Dist. in 1948	Serial Correlation Test[a] 95% level	90% level	Characteristic Root Test[a] 95% level	90% level
1.0	Investment			*CLS*			A	R
2.0	Inventory			neither			R	R
3.4	Production		*R*	*CLI*			R	*R*
4.0	Labor			*CLS*			R	*R*
4.2	Labor			*CLI*	*R*	*R*	*R*	*R*
5.1	Wage		*R*	neither			*R*	*R*
6.2	Consumption	*R*	*R*	neither	*R*	*R*	R	*R*
7.0	Owned housing			neither	*R*	*R*	*R*	*R*
9.0	Rent			*CLS*	*R*	*R*	*R*	*R*
10.0	Rental housing			*CLS*			A	A
11.0	Interest			neither			—	—

Source: Tables 2 and 4, and Section 11, parts (a) and (d).
[a] R means reject; a blank space means accept.

With the exception of the consumption function 6.2,[69] all the equations estimated by the limited information method fit the post-sample year 1948 just as well as they fit the data of the sample period. This is shown by the SETI test.

The predictions for 1948 made by the equations of the reduced form are, on the average over all equations, no better (measured by whether their errors are smaller in absolute value) than predictions made by naive models which simply extrapolate either the value of each variable from the preceding year or the trend between the two preceding years (Table 3).[70]

Note 67 concluded:

(4) (CLI) $W_1 = -8.29 \quad (pX - \mathcal{E}) + 8.95 \quad (pX - \mathcal{E})_{-1} + 3.49t + 15.17$

Note the similarity between 1 and 2 and (except for sign) between 3 and 4. (This sign difference is not an error in computation; it is due to a change in sign of a determinant entering the estimate of the parameter on which the estimates are normalized.)

[68] The equations are given in Sections 5 and 7, and more compactly in Table 2.

[69] And, if P and γ are both relaxed to 0.75, the production function 3.4 and the wage adjustment equation 5.1.

[70] In fact they are worse if the restricted least squares method is used instead of the ordinary least squares method.

However, the reduced form predictions are quite consistently better than the predictions of naive model I for variables that changed more than usual in 1948. Further, the equations of the reduced form may be preferable to naive models for predicting effects of exogenous changes even when both methods make equally large errors in the ordinary prediction of the magnitudes of economic variables, especially when the exogenous changes are unusually large.

The least squares method yields on the average smaller calculated disturbances for 1948 than do our asymptotically superior methods, for both structural and reduced form equations.[71] This is seen by a simple pairwise comparison of calculated disturbances in Tables 2 and 3.

Four equations, as estimated by the limited information method, are rejected by the two-sided test for serial correlation of disturbances, at either the 95 or the 90 per cent level of significance.

Four equations are rejected at the 95 per cent significance level by the characteristic root test of the totality of *a priori* restrictions imposed on a given equation, and seven at the 90 per cent significance level.

Several avenues of future work suggest themselves on the basis of the experience of this paper.

1) Better use could be made of existing economic theory. That is, equations to be estimated should be consistent with the known properties of the equations of micro-economics. Also, a better theory of economic expectations and of behavior under uncertainty would be useful.

2) Studies of narrower sectors of the economy would probably be fruitful, because it is desirable whenever possible to refine our approximations by using variables and equations that apply to more homogeneous groups of firms or individuals. Furthermore, there are several industries and economic sectors for which data, as well as facts pertaining to the technical and institutional environment, are much more plentiful than for the economy as a whole.

3) Cross-section data, i.e., data pertaining to different parts of the economy as of a given point in time such as are obtained in surveys, are becoming increasingly available. It will be possible to combine time series and cross section studies to advantage.

4) One misfortune of the econometrician is that exogenous variables do not vary enough to give him a good idea of their respective influences. The war years are very valuable in this regard, because exogenous changes are ordinarily much larger than in peacetime. Therefore they might be included

[71] This is something to be expected if there are no important changes of structure, and is not contrary to the claims made for limited information estimation; see Section 11, part (c).

in the sample, of course together with appropriate changes in certain parts of the model to allow it to accommodate the wartime government policies.
5) The use of quarterly data would multiply the effective sample size by approximately four,[72] thus producing more accurate estimates, provided the problem of serial correlation can be solved (see next item).
6) The development of practical methods of estimation that do not require the assumption of zero serial correlation of disturbances would be useful. As already mentioned, Chernoff and Rubin have worked on this problem but as yet no attempt has been made to use their results.
7) Mathematical (or experimental)[73] studies to determine the size of the small-sample bias in the estimation of structural parameters by the least-squares method and by the various maximum likelihood methods would be very helpful in deciding which procedure to use.
8) Studies might be made of the effect of estimating the parameters of a model by using data generated by a structure not belonging to the given model, i.e., the effect of estimating from the wrong model. This is a general problem which includes the case of estimating by the least squares method when to do so is not theoretically justified. If a "slightly incorrect" model always or often leads to absurd results, the type of econometrics presented in this paper will suffer a severe setback, because we know from the start that our models are at least slightly incorrect.
9) It would be interesting, though expensive, to estimate the parameters of a fairly large system of equations by the full information maximum likelihood method and analyze the results. But this would not be likely to be immediately useful in getting better estimates unless the sample size were much larger than 22.

Appendix A

Time Series

Until 1942 all time series are as given in Klein (11), pp. 141-3, except that those marked with an asterisk below have been revised as indicated

[72] The effective sample size would be multiplied by *exactly* 4, except for several small points: the fact that one degree of freedom goes into the estimation of each parameter; the possibility of adding four new parameters in order to allow for seasonal changes (this is done by introducing four new exogenous variables x_1, x_2, x_3, x_4, such that in the i^{th} quarter all are 0 except x_i which is 1, and estimating the parameter of each); etc.

[73] Orcutt and Cochrane (18, 19) have used sampling experiments of a type that might be widely applied in getting information of any desired degree of statistical reliability about certain problems that seem to be secure against direct mathematical attack.

in Appendix B, and the X series presented here reflects the correction of a computational error, which has been corrected in Klein (13) also.

	C	I	q	ΔH	D_1	q_1	D_2	D_3	D''
1941						1.348*			
1942		−1.748	1.303	−.379	.870	1.497	.831	.174	1.977
1943		−3.648	1.331	−.436	.514	1.565	.611	.183	1.965
1944	76.833	−3.950	1.351	−1.224	.473	1.621	.539	.143	1.947
1945	80.451	−3.267	1.398	−.425	.462	1.683	.429	.118	1.926
1946	86.517	1.311	1.619	4.020	1.334	2.000	.941	.182	1.926
1947	81.708	2.266	1.964	.345	1.629	2.671	1.162	.172	1.943
1948	82.840	1.889	2.160	3.041	1.858	3.039	1.443	.173	1.969

	G	$Y+T$	Y	p	W_2	W_1	R_1	R_2	r
1939					8.116*	39.959*			
1940					8.348*	43.940*			
1941					9.436*	55.053*			
1942					14.458				
1943					23.224				1.144
1944	70.902	141.769	99.705	1.333	28.211	87.834	10.613	.820	1.146
1945	59.604	135.446	96.286	1.369	29.595	84.908	11.002	.889	1.147
1946	22.835	115.214	89.930	1.561	18.257	91.550	11.790	.978	1.150
1947	17.946	103.285	78.186	1.876	16.008	104.718	12.868	1.105	1.178
1948	17.273	106.548	79.928	2.025	17.529	114.650	14.287	1.162	1.244

	ΔF	v	N'	i	$\mathcal{E}_R$	K	$\mathcal{E}$	X	H
1943				3.16					
1944				3.05				112.299	27.361
1945	571	100.0	31.3	2.87		98.787	7.9	105.142	26.936
1946	1088	100.0	31.8	2.74	976	101.098	9.4	95.339	30.956
1947	1381	100.0	32.7	2.86	853	103.364	10.0	87.304	31.301
1948	1582	100.0	33.6	3.08	888	105.253	10.6	90.263	34.342

	M	N	N_L	w	X
1920	38.464	38.335	38.609	1.080	40.3
1921	38.1005	34.737	39.259	.921	38.8
1922	40.6235	36.335	39.950	.938	42.8
1923	43.249	39.035	40.815	1.040	49.3
1924	46.826	38.744	41.592	1.035	48.5
1925	49.981	39.379	42.044	1.056	52.5
1926	50.876	40.748	43.072	1.087	55.6
1927	53.802	40.792	44.103	1.086	55.8
1928	55.355	40.969	45.128	1.118	56.0
1929	54.555	42.489	46.247	1.132	58.1
1930	53.248	40.397	46.757	1.067	52.3
1931	47.861	37.214	47.313	.951	44.1
1932	44.854	33.816	47.967	.787	35.1
1933	41.532	33.770	48.627	.728	36.7
1934	46.270	36.177	49.127	.788	42.2
1935	51.273	37.162	49.583	.840	47.1
1936	56.360	39.142	49.961	.902	55.3
1937	55.815	41.026	50.433	.999	57.5
1938	58.066	38.657	50.908	.957	52.6
1939	63.253	40.014	51.437	.999	61.0
1940	70.008	41.851	51.722	1.050	66.9
1941	76.336	45.369	51.653	1.213	79.8
1942		47.678	51.427		
1943		48.149	49.821		
1944	130.225	47.111	48.900	1.864	112.299
1945	150.793	45.662	48.181	1.859	105.142
1946	164.004	48.533	52.145	1.886	95.339
1947	170.010	51.019	54.937	2.053	87.304
1948	168.700	52.066	56.021	2.202	90.263

Appendix B

Sources of Data and Construction of Time Series

Construction of time series for 1942 and later, and for the few of Klein's figures for years before 1942 that were revised, is indicated below. My time series are intended to be as consistent as possible with Klein's, since they are extensions of Klein's. The variables denoted by numbers in parentheses correspond to those in the appendices to Klein (11, 13), with the exception of my numbers (13), (14), (15), and (38). The following abbreviations are used:

BAE: Bureau of Agricultural Economics
BLS: Bureau of Labor Statistics
C.C.M.: *Construction and Construction Materials*
F.R.B.: *Federal Reserve Bulletin*
H.L.S.: *Handbook of Labor Statistics*
Klein: L. R. Klein, The Use of Econometric Models as a Guide to Economic Policy, *Econometrica,* 15 (1947), pp. 111-51.
M.L.R.: *Monthly Labor Review*
S.A.U.S.: *Statistical Abstract of the United States*
S.C.B.: *Survey of Current Business*

C: consumption, in billions of 1934 dollars.

$$C = \frac{(1) + (2)}{(3)}$$

(1) = consumer expenditures, Department of Commerce old series, *S.A.U.S.*, 1947, p. 273, for years through 1946. 1947-48 values were obtained from a regression (1939-46) of the old series on the new series (*S.C.B.*, July 1948, p. 16, Table 2; and July 1949, p. 10, Table 2).

(2) = imputed rents on owner-occupied residences, *S.C.B.*, July 1948, p. 24, Table 30; and July 1949, p. 23, Table 30.

(3) = price index of consumption goods, 1934: 1.00, weighted average of the BLS consumers' price index for moderate-income families in large cities (*M.L.R.*, Table D-1) and the BAE index of prices paid by farmers for living (*S.A.U.S.*, 1948, p. 642, and *S.C.B.*, Feb. 1949, p. S-4). The weights are proportional to the urban and rural populations, respectively[1] (*S.A.U.S.*, 1946, p. 14; 1947, p. 15; and 1948, p. 15; and U. S. Bureau of Census, *Current Population Reports,* Series P-20, No. 22, p. 6).

[1] Weights used were, respectively: 1944, .596 and .404; 1945, .586 and .414; 1946, .600 and .400; 1947, .590 and .410; 1948, .584 and .416.

I: net investment in private producers' plant and equipment, in billions of 1934 dollars.

$$I = \frac{(4)}{(5)} + \frac{(6)}{(7)} - \frac{(8)}{(9)} - \frac{(10)}{(7)}$$

(4) = gross expenditures on private producers' nonagricultural plant and equipment, *S.C.B.*, March 1948, p. 24; and February 1949, back cover.

(5) = price index of business capital goods, 1934: 1.00, regression on Solomon Fabricant's index (*Capital Consumption and Adjustment,* pp. 178-9, and private correspondence) of a weighted average of the Aberthaw index (*S.A.U.S.*, 1948, p. 792, and *S.C.B.*, Feb. 1949, p. S-6), the American Appraisal Co. index (*S.A.U.S.*, 1948, p. 792, and *C.C.M.*, May 1949, p. 54), and the BLS index for metals and metal products (*S.A.U.S.*, 1948, p. 296, *M.L.R.*, March 1949, p. 381); weights α, β, and γ, respectively, are such that the weighted average is the same as Fabricant's index in 1934, in 1941, and on the average for 1934-41.

(6) = gross expenditures on farm service buildings and machinery; equal to expenditures on farm buildings excluding operators' dwellings, farm machinery excluding motor vehicles, farm trucks, and farm autos used in production (assumed to be 50 per cent of expenditures on autos in 1942-45 and 40 per cent thereafter), BAE, private correspondence.

(7) = price index of farm capital goods, 1934: 1.00, weighted average of price indexes for building materials for other than housing (*Agricultural Statistics,* 1947, p. 524, and BAE, private correspondence), farm machinery (same), and motor vehicles (BLS metals and metal products index;[2] see (5) above). The weights are proportional to expenditures on each of the three categories of capital goods, respectively. (Because of an error, current dollar expenditures were used as weights instead of constant dollar expenditures, but as the resulting error in I is less than 1 per cent in all cases, and less than 0.1 per cent in most cases, no recomputation was made.)

(8) = depreciation charges on private producers' nonagricultural plant and equipment, regression (1929-43) of Mosak's nonagricultural depreciation (*Econometrica,* 13, 1945, p. 46) on the Department of Commerce depreciation series (*S.C.B.*, July 1947 Supplement, p. 20, Table 4; July 1948, p. 17, Table 4; and July 1949, p. 11, Table 4).

[2] The *BAE* index of motor vehicle prices was discontinued at the start of the war.

(9) = price index underlying depreciation charges, 1934: 1.00, regression (1934-41) of Fabricant's depreciation price index (Fabricant, *Capital Consumption and Adjustment,* p. 183, and private correspondence) on (5).

(10) = depreciation charges on farm service buildings and machinery, BAE, private correspondence.

q: price index of private investment goods, 1934: 1.00.

$$q = \frac{(5) \times (4) + (7) \times (6)}{(4) + (6)}$$

ΔH*:* value of the change in inventories, in billions of 1934 dollars.

$$\Delta H = \frac{(11)}{(12)}$$

(11) = value of change in inventories, Department of Commerce old series, *S.A.U.S.,* 1947, p. 273, and *S.C.B.,* May 1942, p. 12, for years through 1946. 1947-48 values were obtained from a regression (1939-46) of the old series on the new series (*S.C.B.,* July 1948, p. 16, Table 2, and July 1949, p. 10, Table 2).

(12) = BLS wholesale price index of all commodities, 1934:1.00, *F.R.B.,* March 1949, p. 297.

D_1: gross construction expenditures on permanent, owner-occupied, single family, nonfarm residences, in billions of 1934 dollars.[3]

$$D_1 = \frac{1.076 \times 1.126\,[0.63 \times (13) \times (14) + .32 \times (15)]}{(16)}$$

(13) = ratio of *1-family* permanent nonfarm residences started to *total* permanent nonfarm units started, *H.L.S.,* 1947, p. 193; and *M.L.R.,* February 1949, p. 179 (graph), and May 1949, p. 620.

(14) = gross private construction expenditures for new permanent nonfarm residences, *H.L.S.,* 1947, pp. 170-1, and *C.C.M.,* May 1949, p. 6.

(15) = private repairs and maintenance expenditures on nonfarm residences, *C.C.M.,* May 1948, p. 15. (This figure is not available after 1944; hence total residential repairs and maintenance was multiplied by the ratio of *nonfarm* to *total* new residential construction to get an approximation; *C.C.M.,* May 1949, pp. 6, 15.)

1.076 = ratio of average permit valuation of *single-family* urban units to *all* urban units in 1942, *BLS Bulletin 786,* The Construction Industry in the U. S., p. 21, Table 11.

[3] 0.63 = fraction of single-family, nonfarm dwelling units constructed 1935-40 that were owner-occupied in 1940, *Census of Housing, 1940,* III, Part I, Table A-4 (quoted by Klein, p. 144).

1.126 = ratio of average rental value of *owner-occupied* single-family nonfarm residences (constructed 1935-40) to that of *all* single-family nonfarm residences (constructed 1935-40), *Census of Housing, 1940,* III, Part I, Table A-4 (Klein, p. 144).

.32 = ratio of *owner-occupied single-family* nonfarm units to *total* nonfarm units in 1940 (Klein, p. 144).

(16) = American Appraisal Co. index of construction costs (national average), 1934: 1.00, *S.A.U.S.*, 1948, p. 792, and *C.C.M.*, May 1949, p. 54.

q_1: construction cost index, 1934: 1.00.

$$q_1 = (16)$$

D_2: gross construction expenditures on rented nonfarm residences, in billions of 1934 dollars.

$$D_2 = \frac{(17)}{(16)} - D_1$$

(17) = (14) + (15).

D_3: gross construction expenditures on farm residences, in billions of 1934 dollars.

$$D_3 = \frac{(18)}{(19)}$$

(18) = gross construction expenditures on farm residences, *C.C.M.*, May 1948, pp. 8, 15; and May 1949, pp. 8, 15.

(19) = BAE index of farm dwelling construction costs, 1934: 1.00, *C.C.M.*, May 1948, p. 56, and May 1949, p. 58.

D'': depreciation of all residences (farm and nonfarm), in billions of 1934 dollars (on the basis of 3 per cent per year).

$$D'' = (67.6)(.97)^{t-1934}(.03) + \sum_{i=1934}^{t-1} (D_1 + D_2 + D_3)_i$$

$$(.985)(.97)^{t-1-1934}(.03) + (D_1 + D_2 + D_3)_t\,(.015) \qquad \text{for } t > 1934$$

67.6 = estimated value, January 1, 1934, of the stock of residential dwellings in the U. S. (Klein, p. 145).

G: government expenditures for goods and services (not excluding government interest payments) plus net exports plus net investment of nonprofit institutions, in billions of 1934 dollars.

$$G = \frac{(20) - (21) - (22)}{(23)} + \frac{(22)}{(16)} + \frac{(24)}{(12)} + \frac{(25) - 0.1}{(16)}$$

(20) = government expenditures for goods and services, Department of Commerce old series, *S.A.U.S.*, 1947, p. 273, for years through 1946. 1947-48 values were obtained from a regression (1939-46)

of the old series on the new series (*S.C.B.*, July 1948, p. 16, Table 2; and July 1949, p. 10, Table 2).

(21) = government interest payments, *S.C.B.*, July 1948, p. 17, Table 4; and July 1949, p. 11, Table 4.

(22) = public construction expenditures (including work-relief construction), *S.C.B.*, July 1948, p. 25, Table 31; and July 1949, p. 24, Table 31.

(23) = BLS wholesale price index of nonfarm products, 1934: 1.00, *H.L.S.*, 1947, p. 126, and *M.L.R.*, March 1949, p. 381.

(24) = net exports of goods and services and gold, equal to net foreign investment, *S.C.B.*, July 1948, p. 16, Table 1; and July 1949, p. 10, Table 1.

(25) = gross construction expenditures by nonprofit institutions, *S.C.B.*, July 1947 Supplement, p. 44, Table 31; July 1948, p. 25, Table 31; and July 1949, p. 24, Table 31.

0.1 = estimate of depreciation of nonprofit institutions' plant, based on a rate of approximately 3 per cent (Klein, p. 146).

$Y + T$: net national product, in billions of 1934 dollars.

$$Y + T = C + I + \Delta H + D_1 + D_2 + D_3 - D'' + G$$

Y: disposable income, in billions of 1934 dollars.

$$Y = \frac{1}{(3)}\Big[(1) + (2) + (4) + (6) - \frac{(8)}{(9)}(5) - (10) + (11) + (17) + (18) - (16)D'' + (20) + (24) + (25) - 0.1 - (26) - (27) - (28) + (29)\Big]$$

(26) = federal government receipts, *S.C.B.*, July 1948, p. 17, Table 8; and July 1949, p. 12, Table 8.

(27) = state and local government receipts, same sources as for (26).

(28) = net corporate savings (undistributed corporate profits after taxes plus corporate inventory valuation adjustment, plus excess of wage accruals over disbursements, *S.C.B.*, July 1948, p. 17, Table 4; and July 1949, p. 11, Table 4).

(29) = government transfer payments, same sources as for (28).

p: price index of output as a whole, 1934: 1.00.

$$p = \frac{1}{Y + T}\Big[(1) + (2) + (4) + (6) - \frac{(8)}{(9)}(5) - (10) + (11) + (17) + (18) - (16)D'' + (20) + (24) + (25) - 0.1 - (21)\Big]$$

W_2: government wage-salary bill, in billions of current dollars.

$$W_2 = (31)$$

(31) = government wages and salaries, including work relief, Department of Commerce old series, *S.A.U.S.*, 1947, p. 269, for years through 1946. 1947-48 values were obtained from a regression (1939-46) of the old series on the new series with adjustments for income in kind to armed forces (*S.C.B.*, July 1948, p. 16, Table 1; and July 1949, p. 10, Table 1).

W_1: private wage-salary bill, in billions of current dollars.

$$W_1 = (30) - (31)$$

(30) = total employee compensation, including work relief, Department of Commerce old series, *S.A.U.S.*, 1947, p. 269, for years through 1946. 1947-48 values were obtained from a regression of the old series on the new series (*S.C.B.*, July 1948, p. 16, Table 1; and July 1949, p. 10, Table 1).

R_1: nonfarm rentals, paid and imputed, in billions of current dollars.

$$R_1 = (32)$$

(32) = sum of owner-occupied and tenant-occupied nonfarm rents, *S.C.B.*, July 1948, p. 24, Table 30; and July 1949, p. 23, Table 30.

R_2: farm rentals, paid and imputed, in billions of constant dollars.

$$R_2 = (33)$$

(33) = farmhouse rentals, same sources as for (32).

r: index of rents, 1934: 1.00.

$$r = (34)$$

(34) = rent component, 1934: 1.00, of *BLS* consumers' price index for moderate-income families in large cities, *S.A.U.S.*, 1948, p. 302, and *M.L.R.*, March 1949, p. 375.

ΔF: thousands of new nonfarm families.

$$\Delta F = (35)$$

(35) = increase in nonfarm families, thousands, *S.A.U.S.*, 1948, p. 46, and Bureau of Census, *Current Population Reports*, Series P-20, No. 21, p. 9. As the number of families is not given as of the same date each year, adjustments were based on linear interpolation between dates given.

v: percentage of nonfarm housing units occupied at the end of the year, assumed equal to 100.

N^s: millions of available nonfarm housing units at the end of the year.

$$N^s = 31.3 + \sum_{i=1946}^{t} (38)_i$$

(38) = millions of nonfarm housing units finished during the year, *M.L.R.*, March 1948, p. 368, for 1946 and 1947. This series has

been discontinued; for 1948 the number of nonfarm units started was used as an approximation (*M.L.R.*, May 1949, p. 620).

31.3 = millions of available nonfarm dwelling units in November 1945, assuming $v = 100$, *S.A.U.S.*, 1947, p. 799.

i: average corporate bond yield.

$$i = (40)$$

(40) = Moody's corporate bond yield, *F.R.B.*, 1947, p. 1519; and 1949, p. 275.

$\mathcal{E}_R$: excess reserves, in millions of current dollars.

$$\mathcal{E}_R = (43)$$

(43) = annual average of monthly figures for excess reserves, *F.R.B.*, 1947, pp. 551, 987, 1377; 1948, pp. 187, 523, 965, 1373; 1949, p. 137.

K: end of year stock of private producers' plant and equipment, in billions of 1934 dollars.

$$K = 107.8 + \sum_{i=1935}^{t} I_i \qquad t \leqq 1945$$

$$K = 1.0 + 107.8 + \sum_{i=1935}^{t} I_i \qquad t \geqq 1946$$

107.8 = end of 1934 stock of private producers' plant and equipment (Klein, p. 148).

1.0 = estimate of surplus property transferred to the private sector at the close of the war (Klein, p. 150).

H: end of year stock of inventories, in billions of 1934 dollars.

$$H = 21.8 + \sum_{i=1935}^{t} (\Delta H)_i$$

21.8 = end of 1934 stock of inventories (Klein, p. 149).

$\mathcal{E}$: excise taxes, in billions of current dollars.

$$\mathcal{E} = (45)$$

(45) = excise taxes, regression (1931-41) of Klein's excise series (p. 149) on the sum of federal excises plus state and local sales and social insurance taxes (*S.C.B.*, July 1947 Supplement, p. 21, Table 8; July 1948, p. 17, Table 8; and July 1949, p. 12, Table 8).

X: private output excluding housing services, in billions of 1934 dollars.

$$X = Y + T - \frac{1}{p}(W_2 + R_1 + R_2)$$

M: end of year money supply, in billions of current dollars.

$$M = (46) \qquad t \leqq 1922$$
$$M = (47) \qquad t \geqq 1923$$

(46) = demand and time deposits adjusted plus currency outside banks, average of June 30 figures before and after, Federal Reserve Board, *Banking and Monetary Statistics,* p. 34.

(47) = demand and time deposits adjusted plus currency outside banks, Dec. 31 figures, *Banking and Monetary Statistics,* pp. 34-5, and *F.R.B.*, 1949, p. 265.

N: labor input, in millions of full time equivalent man-years.

$$N = (48) \qquad t \leqq 1928$$
$$N = (49) \qquad t \geqq 1929$$

(49) = number of full time equivalent persons engaged in production in all private industries, excluding work relief, *S.C.B.*, July 1947 Supplement, p. 40, Table 28; July 1948, p. 23, Table 28; and July 1949, p. 22, Table 28.

(48) = regression (1929-38) of (49) on Kuznets' estimates of total persons engaged in private production (*National Income and Its Composition,* pp. 314-5, 346-7).

N_L: labor force, including work-relief employees but excluding other government employees, in millions of man-years.

$$N_L = (50) - (51)$$

(50) = civilian labor force, Census definition, *H.L.S.*, 1947, p. 36, and *S.C.B.*, February 1949, back cover, for years after 1928. 1920-28 values were obtained from a regression (1929-39) of the Census series on the National Industrial Conference Board series (*Economic Almanac,* 1944-45, p. 43).

(51) = government full time equivalent civilian employees excluding work-relief employees, *S.C.B.*, July 1947 Supplement, p. 36, Table 24; July 1948, p. 22, Table 24; and July 1949, p. 20, Table 24, for years after 1929. 1920-28 values were obtained from a regression (1929-38) of the above series on Kuznets' estimates of the same quantity (*National Income and Its Composition, 1919-1938,* pp. 314-5).

w: private money wage rate, in thousands of current dollars per man-year.

$$w = \frac{W_1}{N}$$

Appendix C

Time Series for 1946-1947

During the discussion at the Conference on Business Cycles Research in November 1949, Lawrence Klein pointed out a discrepancy in the time series for 1946 and 1947 which were used in this paper: the series for real net national product $Y + T$ and for real private output X show decreases of about 10 per cent from 1946 to 1947, while during the same two years the series for private employment N rose about 5 per cent and the Federal Reserve Board index of industrial production rose 10 per cent. Since these four series are meant to measure magnitudes that have to move closely together (except that the agricultural sector is not represented in the Federal Reserve index), it is clear that something is wrong. It is difficult to see how the series for employment and industrial production could be seriously in error for this period, but the series for $Y + T$ and X might be thrown off by either or both of two causes.

First, the series for $Y + T$ and X are constructed by adding component series, each of which is first expressed in current prices, then deflated by an appropriate price index. It is very likely that the published price indexes (which were used in the paper) are too low for the years toward the end of the reign of price controls, including 1946, because of failure to take account of reductions in quality and service, black market activities, and the practice on the part of manufacturers of concentrating their output in their more expensive lines. This understatement has been estimated by the Technical Committee on the consumers' price index (also known as the Mitchell Committee) not to exceed about 4 per cent in any year (see the *Economic Report of the President,* January 1950, pp. 156 and 169), and by various others to be considerably larger. It can be expected to have disappeared by some time in 1947, because virtually all controls were lifted in November 1946, and many had been lifted or relaxed before then. Therefore it is a good surmise that while the published price indexes are too low in 1946, they again measure approximately what we want them to measure in 1947. If this is true, the deflated series for $Y + T$ and X are too high in 1946, and therefore their apparent drop from 1946 to 1947 is partly or wholly illusory — there may even have been a rise, camouflaged by the understated 1946 price indexes. My guess would be that the entire discrepancy is not to be explained in this manner, however.

Second, as indicated in Appendix B, the time series used were extensions of Klein's own time series, based like his on the series released by the Department of Commerce before the publication in 1947 of its revised national income series. Some of the 1947-48 figures were obtained from regressions of the unrevised series on the corresponding revised series. It would have been sounder to adjust all the time series, including Klein's, to conform to the revised Department of Commerce series, or failing that, to obtain the 1947-48 extrapolations of the unrevised series by adjusting the revised series for changes in definition instead of using regressions.

Similar discrepancies, of comparable magnitude, are obtained for

1946-47 for the whole economy and for separate industries if the national income originating in the economy and in each of several industries, deflated by the corresponding wholesale price index, is compared industry-wise with the number of full time equivalent persons engaged in production or with the Federal Reserve index of industrial production. (They are clearly visible, even though the industrial classifications are not quite the same in the Federal Reserve index and national income accounts as in the wholesale price index.) Because these discrepancies are comparable in magnitude to the one pointed out by Klein, it seems likely that it is unnecessary to look to my regression technique for an explanation of the error in the relationship of the 1946 to the 1947 figures; it even seems likely that the regression technique made no significant contribution to that error (though no doubt it introduced others).

It remains to determine the effect of the discrepancies on the results of the paper. Of course the most reliable way would be to revise all the data and re-estimate all the equations. Here it is possible only to try to obtain a rough idea of the effect, by means of some approximation 'corrections' consisting of making changes in some of the 1946-47 time series so that they become consistent with the Federal Reserve index of industrial production, then re-estimating certain of the structural and reduced form equations by the ordinary least squares method. The detailed steps and results of this exploratory 'correction' procedure are explained below.

The time series for real private output X, disposable income Y, and consumption C are accepted as correct for 1947, and are 'corrected' for 1946. Let unprimed symbols stand for the values underlying the original computations of the paper, and primed symbols for the 'corrected' values. Then,

$$X'_{1946} = \frac{170}{187} X_{1947}$$

where $\frac{170}{187}$ is the ratio of the 1946 to 1947 values of the Federal Reserve index of industrial production. (If employment were used as the correction standard instead, a less drastic reduction factor than $\frac{170}{187}$ would result; however, we use $\frac{170}{187}$ so as to be sure not to *under*estimate the effect of the discrepancy we are analyzing.) Also,

$$Y'_{1946} = \frac{X'_{1946}}{X_{1946}} Y_{1946}$$

$$C'_{1946} = \frac{X'_{1946}}{X_{1946}} C_{1946}$$

$$\left[\left(\frac{M}{p}\right)_{-1}\right]'_t = \frac{X'_{1946}}{X_{1946}} \left[\left(\frac{M}{p}\right)_{-1}\right]_t \qquad t = 1946, 1947$$

The purpose of these changes is to gear output X for 1946 to the Federal Reserve index (while accepting X for 1947), and then to make the same percentage change in the 1946 values of C, Y, and $\left(\frac{M}{p}\right)_{-1}$ as was made in the 1946 value of X. The 'corrected' values are shown in Table C1.

Table C1

CORRECTED TIME SERIES

	X	C	Y	$(M/p)_{-1}$
1946	79.367	72.023	74.864	91.695
1947	....	....	....	87.462

Then the production functions (3.2) and (3.4), the consumption function (6.2), and the reduced form equations for C and Y are re-estimated by the ordinary least squares method, incorporating the above changes into the time series. The results of the structural re-estimation are shown as the *CLS'* estimates in Table C2, which reproduces the relevant *CLS* estimates from Table 2 for convenience in comparison.

Table C2

RE-ESTIMATION OF CERTAIN STRUCTURAL EQUATIONS

Eq.	Estimates of Parameters (and of Standard Errors) (1)	(2)	(3)	(4)	(5)	S (8)	Obs. Value 1948 (10)	Calc. Value 1948 (11)	Calc. Dist. 1948 (12)
3.2 Production	X	N	K_{-1}	t	1				
CLS	1	2.820	−.296	.55	−26.50	3.23	90.26	99.02	−8.76
		(.25)	(.16)	(.14)					
CLS'	1	2.771	−.044	.34	−51.8	2.94	90.26	93.73	−3.47
		(.23)	(.15)	(.13)					
3.4 Production	X	N	t	1					
CLS	1	3.074	.43	−68.13		3.42	90.26	99.16	−8.90
		(.22)	(.13)						
CLS'	1	2.809	.32	−58.01		2.87	90.26	93.65	−3.39
		(.18)	(.11)						
6.2 Consumption	C	Y	$(M/p)_{-1}$	t	1				
CLS	1	.583	.297	−.27	7.07	1.57	82.84	75.97	6.87
		(.06)	(.04)	(.10)					
CLS'	1	.614	.329	−.30	3.88	2.20	82.84	77.65	5.19
		(.08)	(.10)	(.18)					

The results of the re-estimation of the reduced form equations are shown opposite the primed variables in Table C3, which reproduces certain parts of Table 3 for convenience in comparison.

Table C3

RE-ESTIMATION OF CERTAIN REDUCED FORM EQUATIONS

Var.	Obs. Value 1948 (1)	Predictions 1948 RF LS (4)	Predictions 1948 Naive Models I (6)	Predictions 1948 Naive Models II (7)	Errors 1948 RF LS (8)	Errors 1948 Naive Models I (10)	Errors 1948 Naive Models II (11)	Naive Model Test Verdicts I (12)	Naive Model Test Verdicts II (13)
C	82.8	80.2	81.7	76.9	2.6	1.1	5.9	*N*	*RF*
C'	82.8	90.3	81.7	91.4	−7.5	1.1	−8.6	*N*	*RF*
Y	79.9	70.9	78.2	66.4	9.0	1.7	13.5	*N*	*RF*
Y'	79.9	86.1	78.2	81.5	−6.2	1.7	−1.6	*N*	*N*

From Table C2 it appears that the differences between the estimates obtained in this paper and the estimates that would be obtained if the time series discrepancies were corrected are not likely to be negligible, and that some of the structural equations would be likely to fit better in 1948 as a result of the corrections. From Table C3 it appears that the correction process would be likely to produce non-negligible changes in the predictions made by the reduced form. But Table C3 does not indicate that *better* predictions of the important variables C and Y would be obtained if the time series discrepancies were corrected.

The variables whose time series are most likely to be changed by a revision of the data, in a way important enough to influence the results, are $C, I, q, \Delta H, G, Y, T, p$, and X. Those likely to be affected in an unimportant way (because they are small or stable) are D_1, q_1, D_2, D_3, D'', r, K, and H. Those not likely to be affected at all (because they are independent of price indexes) are $(pX - \mathcal{E})$, W_1, W_2, R_1, R_2, ΔF, v, N^s, i, $\mathcal{E}_R$, $\mathcal{E}$, M, N, N_L, and w. Accordingly, equations 1.0, 2.0, 3.4, 4.2, and 6.2 are likely to be affected in a significant way because they are dominated by variables from the first of the aforementioned groups. Similarly, equations 4.0, 5.1, 7.0, 9.0, 10.0, and 11.0 are not likely to be affected significantly because they are not dominated by variables from the first group.

Naive model I as applied to 1948 is unaffected by the changes made here because it does not reach as far into the past as 1946. Naive model II is affected, however, and will be led by the changes to make uniformly higher predictions of deflated quantities (usually an improvement in performance) and a lower prediction of the general price index (also an improvement).

The upshot of the calculations based on these approximate 'corrections' is something like this: if the data were revised and the equations re-estimated, the estimates of the parameters would be changed, and the 1948 fit of some structural equations would probably be improved, but there is no evidence that the predictions of important variables by the reduced form would be improved.

Appendix D

Choice of Predetermined Variables for Estimation by the Abbreviated Variant of the Limited Information Method

The limited information estimates of any structural equation depend upon observations of a subset of the predetermined variables that are not in the equation being estimated but are in the system. The elements of this subset are called z^{**}'s and there must be at least as many as $H-1$ of them if H is the number of jointly dependent variables in the equation being estimated (see text, Sec. 1). Of course, there may be more than $H-1$; if so, the estimates will be better. In our model the largest value of $H-1$ for any equation is 4, for equation 3.0; if this is excepted, the largest value is 2, for each of several equations. Therefore the number of z^{**}'s *required* for any equation is 2 except in the case of equation 3.0, which requires 4.

Now there are 25 predetermined variables in the complete model, and no equation contains more than 4. Thus, for each equation there are at least 21 variables available for use as z^{**}'s, and so there is an arbitrary choice of z^{**}'s to be made for each equation. If there were no costs in money and in degrees of freedom, one would always use all the available variables as z^{**}'s. Because of these costs, a proper subset of the available variables has been used in each case, i.e., the abbreviated variant of the limited information method has been used.

The stochastic equations have been divided into four groups in such a way as to minimize the intersection of the set of jointly dependent variables in any group with the corresponding set for any other group; in fact every such intersection is empty. Then for any equation the set of z^{**}'s is the set of *all* predetermined variables in the group to which the equation belongs, minus the set of predetermined variables appearing in the equation (see the accompanying table).

GROUP	EQUATIONS	VARIABLES: Jointly Dependent	VARIABLES: Predetermined
I	(1.0), (2.0), (3.4), (4.0), (4.2), (5.1)	$I, \frac{pX-\mathcal{E}}{q}, H, p, (X-\Delta H), X, N, W_1, (pX-\mathcal{E}), w/p, w, \Delta p, (N_L-N)$	$\left(\frac{pX-\mathcal{E}}{q}\right)_{-1}, K_{-1}, H_{-1}, t, (pX-\mathcal{E})_{-1}, w_{-1}, (N_L-N)_{-1}, p_{-1}$
II	(6.2), (7.0), (9.0)	$C, Y, D_1, r/q_1, (Y+Y_{-1}+Y_{-2}), \Delta r$	$(M/p)_{-1}, t\ \Delta F, v_{-1}, 1/r_{-1}$
III	(10.0)	D_2, i	$r_{-1}, (q_1)_{-1}, (q_1)_{-2}, \Delta F_{-1}, i_{-1}, t$
IV	(11.0)	Δi	$i_{-1}, \mathcal{E}_R, t$

Klein's grouping of equations was quite similar. In particular for group I he used exactly the same predetermined variables as I did, except that in place of w_{-1} and $(N_L - N)_{-1}$ he used H_{-2} and X_{-1}. This is mentioned here because of its possible bearing on certain anomalies in the *CLI* estimates of equations 1.0 and 4.0 of group I. The matter is discussed in the text in Section 11, part (f).

Appendix E

Estimation of the Parameters of the Reduced Form

This appendix is a note on the restricted least-squares method of estimating reduced form parameters, referred to in Section 4. We first describe the method assuming that a one-element subset of structural equations is chosen to provide the restrictions.

Suppose there is a model consisting of G equations in G jointly dependent variables y and K predetermined variables z. Suppose that one of its equations is

$$(1) \qquad \beta_1 y_1 + \cdots + \beta_H y_H + 0 + \cdots + 0 + \gamma_1 z_1 + \cdots + \gamma_{K^*} z_{K^*} + 0 + \cdots + 0 = u$$

where $H < G$ and $K^* < K$. Consider H equations of the reduced form,

$$(2) \qquad y_i = \sum_{1}^{K^*} \pi_{ik} z_k + \sum_{K^*+1}^{K^*+K^{**}} \pi_{ik} z_k + v_i \qquad i = 1, \cdots, H$$

where K^{**} is the number of predetermined variables assumed to be known to be in the model but not in 1. Then $K^{**} \leqq K - K^*$. The parameters π_{ik} can be estimated by least-squares. The least-squares estimates can be made more efficient by altering them to take account of the restrictions implied by the zeros in 1, as follows. It must be possible to get equation 1 from a linear combination of equations 2, in fact, from that combination obtained by taking β_i times the i^{th} equation of 2, $i = 1, \cdots, H$, and summing the results. This means that there are K^{**} equations, one for each z_k excluded from 1, thus:

$$(3) \qquad \sum_{i=1}^{H} \beta_i \pi_{ik} = 0 \qquad k = K^* + 1, \cdots, K^* + K^{**}$$

Now if $K^{**} > H - 1$, i.e., if 1 is overidentified, 3 is overdetermined. Hence if 3 is to hold, and it must, a restriction is implied on the matrix of the π_{ik}, $i = 1, \cdots, H$, $k = K^* + 1, \cdots, K^* + K^{**}$, keeping its rank down to $H - 1$. This restriction may be applied to the matrix of least-squares estimates of the π_{ik}, to make them conform to the restrictions implied by the zeros in 1. The computation is not difficult, once the limited information estimates for 1 are obtained.

Similarly, if there are other structural equations besides 1 which also contain some one of the jointly dependent variables $y_1, \cdots, y_H$, say y_1, the estimates of the parameters of the reduced-form equation for y_1 can be made to conform simultaneously to the restrictions implied in the form of two, three, ···, or all these other structural equations as well. This further increases the efficiency of the estimates, but makes them more difficult to compute.

Appendix F

CALCULATED DISTURBANCES FOR CLI LIMITED INFORMATION ESTIMATES OF EQUATIONS

	1.0	2.0	3.4	4.0	4.2	5.1	6.2	7.0	9.0	10.0	11.0
1921										.03	.37
1922	2.11	−.26	3.04	37.07	−1.08	.033	−.61	.15	−.010	−.01	−.66
1923	3.93	1.05	1.19	97.88	.09	.019	−.89	.18	−.014	.06	.18
1924	6.38	−.38	.76	−6.20	.28	−.017	.37	.38	.016	.10	−.09
1925	−2.56	.12	2.43	59.97	−.93	−.033	.05	.44	−.004	.18	−.24
1926	−2.60	.73	1.06	34.73	−.41	−.003	.83	.20	−.002	.19	−.25
1927	.56	.25	.65	−18.03	.03	.005	1.54	.19	.010	.06	−.29
1928	−3.41	−.23	−.14	−.51	.59	.009	2.18	.28	.018	.02	−.13
1929	−3.55	1.21	−2.96	10.87	1.54	.003	1.33	−.16	.034	.34	.19
1930	2.45	.64	−3.13	−98.53	1.68	−.016	−.11	−.25	.041	−.75	−.08
1931	7.87	−.53	−2.52	−149.14	2.07	.012	−2.25	−.36	.013	−.28	.76
1932	12.19	−1.18	−2.09	−141.32	1.41	−.028	−3.27	−.47	−.050	−.21	1.38
1933	2.35	−.43	−.83	−20.27	−.12	−.036	−1.24	−.58	−.079	−.01	−.25
1934	−1.11	−.63	−2.83	44.62	−.09	−.043	−.48	−.41	−.019	.004	−.40
1935	−2.36	−.83	−1.29	28.41	−.88	−.022	−.09	−.13	.009	.13	−.17
1936	−5.34	.53	.66	42.52	−1.40	.006	.47	−.04	.019	−.004	−.40
1937	1.63	.09	−3.11	9.26	.64	.003	.80	−.07	.039	.02	−.03
1938	8.27	−1.33	−1.58	−108.97	.56	−.004	2.08	.12	.029	−.06	.32
1939	−6.04	−.28	2.38	27.75	−.73	.025	1.01	.21	.006	.30	−.06
1940	−4.85	.49	2.45	12.84	−.76	.002	.34	.38	.010	.17	.20
1941	−17.27	1.49	4.61	115.82	−1.25	.036	−2.22	.58	.019	.04	.08
1946	1.48	.71	8.53	−38.12	−2.26	−.033	−1.55	−.46	−.050	.15	−.33
1947	1.70	−1.36	−7.24	60.24	1.25	.013	1.88	−.16	−.032	−.22	−.10
1948	−8.37	1.19	−7.82	90.56	.77	.17	6.88	.03	.011	−.19	.07

Appendix G

REFERENCES

1) Anderson, T. W., and H. Rubin, Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *Annals of Mathematical Statistics,* 20 (1949), pp. 46-63, and The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *Annals of Mathematical Statistics,* 21 (1950), pp. 570-82.
2) Bronfenbrenner, J., 'Extent of Least Squares Bias in Estimating a Single Stochastic Equation in a Complete System', Cowles Commission Discussion Paper (unpublished), 1949.
3) Eisenhart, C., M. Hastay and W. A. Wallis (editors), *Techniques of Statistical Analysis* (McGraw-Hill, 1947), Ch. 2.

4) Frisch, Ragnar, Propagation Problems and Impulse Problems in Dynamic Economics, *Economic Essays in Honour of Gustav Cassel* (London, Allen & Unwin, 1933).
5) Haavelmo, T., The Statistical Implications of a System of Simultaneous Equations, *Econometrica,* 11 (1943), pp. 1-13.
6) Haavelmo, T., The Probability Approach in Econometrics, *Econometrica,* 12 (1944), Supplement.
7) Haavelmo, T., Methods of Measuring the Marginal Propensity to Consume, *Journal of the American Statistical Association,* 42 (1947), pp. 105-22.
8) Hart, B. I., and John von Neumann, Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance, *Annals of Mathematical Statistics,* 13 (1942), pp. 207-14.
9) Hotelling, Harold, Problems of Prediction ('The Prediction of Personal Adjustment': A Symposium), *American Journal of Sociology,* 48 (1942-43), pp. 61-76.
10) Kalecki, M., *Essays in the Theory of Economic Fluctuations* (London, Allen & Unwin, 1939).
11) Klein, L. R., The Use of Econometric Models as a Guide to Economic Policy, *Econometrica,* 15 (1947), pp. 111-51.
12) Klein, L. R., Theories of Effective Demand and Employment, *Journal of Political Economy,* 50 (1947), pp. 108-31.
13) Klein, L. R., *Economic Fluctuations in the United States, 1921-1941,* Cowles Commission Monograph 11 (Wiley, 1950).
14) Koopmans, T., Identification Problems in Economic Model Construction, *Econometrica,* 17 (1949), pp. 125-44.
15) Koopmans, T., Statistical Methods of Measuring Economic Relations, Cowles Commission Discussion Paper (unpublished), 1948.
16) Koopmans, T. (editor), *Statistical Inference in Dynamic Economic Models,* Cowles Commission Monograph 10 (Wiley, 1950).
17) Marshall, A., A Test of Klein's Model III for Changes of Structure, to be published.
18) Orcutt, G., and D. Cochrane, Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms, *Journal of the American Statistical Association,* 44 (1949), pp. 32-61.
19) Orcutt, G., and D. Cochrane, A Sampling Study of the Merits of Auto-regressive and Reduced Form Transformations in Regression Analysis, *Journal of the American Statistical Association,* 44 (1949), pp. 356-72.
20) Rubin, H., 'Systems of Linear Stochastic Equations', Cowles Commission Discussion Paper (unpublished), 1947.

21) Rubin, H., 'The Approximate Distribution of Calculated Disturbances', Cowles Commission Discussion Paper (unpublished), 1948.
22) Samuelson, P. A., *Foundations of Economic Analysis* (Harvard University Press, 1947).
23) Tinbergen, Jan, *Statistical Testing of Business Cycle Theories, II: Business Cycles in the United States of America, 1919-1932* (League of Nations, Economic Intelligence Service, Geneva, 1939).
24) Wald, A., and J. Wolfowitz, Tolerance Limits for a Normal Distribution, *Annals of Mathematical Statistics,* 17 (1946), pp. 208-15.

COMMENT

MILTON FRIEDMAN, *University of Chicago*

First, may I congratulate Carl Christ and the Cowles Commission for undertaking to test the predictive value of Klein's econometric model and for the thoroughly objective and scientific manner in which they have performed this task. Economics badly needs work of this kind. It is one of our chief defects that we place all too much emphasis on the derivation of hypotheses and all too little on testing their validity. This distortion of emphasis is frequently unavoidable, resulting from the absence of widely accepted and objective criteria for testing the validity of hypotheses in the social sciences. But this is not the whole story. Because we cannot adequately test the validity of many hypotheses, we have fallen into the habit of not trying to test the validity of hypotheses even when we can do so. We examine evidence, reach a conclusion, set it forth, and rest content, neither asking ourselves what evidence might contradict our hypothesis nor seeking to find out whether it does. Christ and the Cowles Commission have not followed this easy path. They have revised the parts of Klein's econometric model that fit postwar experience least well, then, and this is the important step, have used the revised model to predict an additional year and compared the results with what actually happened. The fact that the results suggest that Klein's experiment was unsuccessful is in some ways less important than the example they set the rest of us to go and do likewise. After all, most experiments are destined to be unsuccessful; the tragic thing is that in economics we so seldom find out that they are.

Klein's model is not the only attempt to construct a system of simultaneous equations to predict short-time changes in important economic phenomena. Probably the most ambitious was the model constructed by Jan Tinbergen on the basis of United States data for 1919-32. Slightly different in character but in the same general class were the equations computed under the direction of Gardiner C. Means at the National Resources Board and published in *Patterns of Resource Use* (1939). More recently, Colin Clark published in *Econometrica* (April 1949) a rather simpler model for the United States economy. And there are still others, both published and unpublished.

These systems of equations all describe adequately the data from which they were derived; that is, they yield high correlation coefficients and most of the estimated parameters are several times their standard errors. But, as is fairly widely recognized by now, the fact that the equations fit the data from which they were derived is a test primarily of the skill and patience of the analyst; it is not a test of the validity of the equations for any broader body of data.[1] Such a test is provided solely by the consistency of the equations with data not used in their derivation, such as data for periods subsequent to the period analyzed. It is my impression that this test has at various times been applied to Tinbergen's and Means' equations and that neither survived it satisfactorily. Colin Clark's model, as far as I know, has not yet been tested.

Christ accepted tests by Andrew Marshall as a basis for revising Klein's original model. He then proceeded to test his revision of Klein's model in two main ways: first, he tested the internal consistency of the equations by seeing whether errors in predicting 1948 values were larger than might be expected on the basis of the unexplained variation in the sample years; second, he compared the predictions of his econometric model with the predictions of what he terms "naive" models.

As the tests of internal consistency seem to me far less important than the naive model tests, I shall add little to what Christ says about them. I wish to mention only that the choice of the probabilities used in defining the tolerance interval (γ and P in Christ's notation), though not discussed by Christ, is critical. The choice of sufficiently high values for these probabilities will assure the acceptance of almost any equation, no matter how bad the prediction, though of course only at the expense of a high chance of failing to reject the equation when it is 'false'. It is my hunch that, given the size of sample, the values initially chosen by Marshall and used by Christ involve an unduly small risk of rejecting a 'correct' equation compared with the risk of accepting an 'incorrect' equation.

The naive model tests deserve somewhat more analysis. The naive

[1] See my review of J. Tinbergen, Business Cycles in the United States of America, 1919-1932, *American Economic Review*, September 1940.

models should not be taken seriously as techniques for actually making predictions; they are not competing theories of short-time change. Their function is quite different. It is to provide a standard of comparison, to set the zero point, as it were, on the yardstick of comparison. We say that the appropriate test of the validity of a hypothesis is the adequacy with which it predicts data not used in deriving it. But how shall we assess the adequacy of prediction? Obviously we need not require perfect prediction; so the question is when are the errors sufficiently small to regard the predictions as unsuccessful? We cannot judge by the absolute size of the error; on what grounds are we to say that an error of, say, \$1 billion is either small or large? Nor do percentage errors help much, even though they seem intuitively more relevant. An error of 2 per cent means one thing if the variable being predicted never varies by more than 3 per cent and quite a different thing if it usually varies by 50 per cent. Moreover, the percentage error is itself really arbitrary. For example, suppose we know income and seek to predict savings and consumption expenditures. Since consumption expenditures will be something like 10 times as large as savings, a 20 per cent error in savings will be approximately a 2 per cent error in consumption. Which is the appropriate number for judging the adequacy of the prediction? The 2 per cent or the 20 per cent error?

If predictions are made for several years (or other units) one simple method of testing the accuracy of the predictions is by the correlation between the predicted and the actual values. This can be computed and compared with the correlations to be expected between chance series, and the prediction judged a success if the correlation is higher than might reasonably be expected from chance alone. But this test, too, has its defects: it is likely to be relatively insensitive for a small number of predicted values; it may require an estimate of the serial correlation among observations if the appropriate sampling distribution is to be used; it is not clear what the appropriate alternative hypotheses are in terms of which the test of significance should be chosen.

The naive models provide an alternative, though related, standard of comparison, which can be used for one year or many years, and which takes account of serial correlation. They are in some sense the 'natural' alternative hypotheses – or 'null' hypotheses – against which to test the hypothesis that the econometric model makes good predictions. The reason can be easily seen. The essential objective behind the derivation of econometric models is to construct an hypothesis of economic change; any econometric model implicitly contains a theory of economic change. Now given the existence of economic change, the crucial question is whether the theory implicit in the econometric model abstracts any of the essential forces responsible for the economic changes that actually occur. Is it better, that is, than a theory that says there are no forces making for change? Now naive model I, which says the value of each variable next year will be identical with its value this year, is precisely such a theory; it denies, as it were, the existence of any forces making for changes from one year to the next. In the language of the econometric model, it says that the appropriate structure is one in which all the equations contain only a constant term, the rest of the parameters being zero. If the econometric model does no better than this naive model, the implication is that it does not abstract any of the essential forces making for change; that it is of zero value as a theory explaining year to year change.

Of course, there are many varieties of change, and many different objectives may be set for an econometric model or for any other theory of change. The forces that are essential in explaining changes from one year to the next may not be the same as those that are essential in explaining changes over a two-, or a five-, or a twenty-year period. And for each of these types of change there is an appropriate naive model of type I. The fact that an econometric model is rejected for one class of change does not mean that it will be rejected for another; but neither, of course, is there any reason to believe that it will not be.

Change can be differentiated also by criteria other than the period considered. In particular, we frequently distinguish between what is called 'secular' and 'cyclical' change. This is the role of naive model II, which says that the value of each variable next year will differ from its value this year in the same direction and by the same amount as its value this year differed from its value last year. This is a theory of 'pure' secular change, as it were; and it seems to me appropriate if and only if the model being tested has passed naive model test I satisfactorily. In that case, the implication is that the model has successfully abstracted some essential forces making for change, and the question can then be asked whether it has isolated secular forces alone or cyclical forces as well.

Christ's revision of Klein's model does no better than naive model I for the one year for which Christ could make the test, 1948. The econometric model makes larger errors than the naive model for approximately half the variables predicted, and its average error is, if anything, larger than the average error of the naive model.

One is tempted to add that the test is biased in favor of the econometric model because of the way exogenous variables are treated. Christ used the actual values of the exogenous variables for 1948 whereas in making a prediction for a future year it would be necessary to predict the exogenous variables independently. But this is not a valid objection; Christ's procedure is the correct one. The model claims to make only conditional predictions: if the exogenous variables are such and such, the endogenous variables will be such and such. And the important first question is whether it can make such conditional predictions.

Of course, one swallow does not make a spring; and one must be careful of generalizing too broadly from tests based on predictions for one year. Perhaps if the model were tested for additional years the unfavorable verdict would be reversed; all one can say is that the evidence so far assembled contradicts the hypothesis of short-time economic change implicit in the econometric model. It is highly desirable that additional evidence be accumulated; but meanwhile I shall proceed on the assumption that additional evidence would not reverse Christ's tentative conclusion.

Christ suggests one qualification to the conclusion that, on the basis of existing evidence, the particular econometric model he tested is worthless. He writes, "Even if such a naive model does predict about as well as our econometric model, our model may still be preferable because it may be able to predict consequences of alternative policy measures and of other exogenous changes, while the naive model cannot." But this argument is at best misleading, at worst invalid. The naive model can make such predictions too: one can simply assert that a proposed change in policy or in an exogenous variable will have no effect. If this kind of prediction worked as well as the econometric predictions for a change from one year to the next, might it not work as well for policy changes also? Note that the evidence implicitly used in predicting the effect of policy changes by means of the econometric model is derived from year to year changes in the basic data, i.e., from precisely the kind of changes the naive model test suggests the econometric model is incompetent to predict. To put the point in another way, the assertion that the econometric model can be used to predict the consequences of policy changes implicitly assumes that the theory of change implicit in the econometric model abstracts some of the essential forces determining economic change; stated loosely, that the model is an approximation to the 'correct' one, and that the parameters are better estimated by giving them the values obtained from the estimated econometric structure than by setting them equal to zero. Now it is precisely these propositions that the naive model test contradicts.

Of course, the policy changes to be predicted may differ in character from the year to year changes that the econometric model failed to predict; and, as Christ suggests, the model may predict the one kind of change even though it does not predict the other. But then it is a pure act of faith to assert that the econometric model can predict the effect of policy changes, and there is no reason for anyone else to share this faith until some evidence for it is presented. Surely, the fact that the model fails to predict one kind of change is reason to have less rather than more faith in its ability to predict a related kind of change.

Granted that this particular experiment in constructing an econometric model must be judged a failure on the basis of present evidence, what implications does this have for future work? One possibility, already mentioned, is that this failure is a freak; that further evidence will show that this model can predict successfully and that one should await such further evidence. Another possibility is that the defects of this model are peculiar to it and not to econometric models of this general kind; that examination of the economic theory implicit in this model, of the detailed shortcomings of individual equations and the like, will permit the construction of an improved model along the same general lines that will work successfully. Neither possibility can be categorically rejected. Like any other prediction, the assertion that it will or will not be realized is a prediction that cannot be made with certainty. My own hunch, however, is that neither possibility will be realized; that additional evidence on this particular model will strengthen rather than reverse the conclusion suggested by the existing evidence and that attempts to proceed now to the construction of additional models along the same general lines will, in due time, be judged failures.

In part, this hunch is simply an extrapolation of experience: as already noted, Klein's model is by no means the first of its general type that has been constructed and tested and so far none has survived the test of ability to predict. But this empirical extrapolation is by itself unsatisfactory. The fundamental premise underlying work in this field is that there is order in the processes of economic change, that sooner or later we shall develop a theory of economic change that does abstract essential elements in the process and does yield valid predictions. When and if such a theory is developed, it will clearly be possible to express it in the form of a system of simultaneous equations of the kind used in the econometric model – mathematics is after all a rather flexible and highly useful language into which practically any economic theory can be translated. Does it not then follow that despite the unsatisfactory results to date, the appropriate procedure is to continue trying one after another of such systems until one that works is discovered?

I think the answer is no. Granted that the final result will be capable of being expressed in the form of a system of simultaneous equations applying to the economy as a whole, it does not follow that the best way to get to that final result is by seeking to set such a system down now. As I am sure those who have tried to do so will agree, we now know so little about the dynamic mechanisms at work that there is enormous arbitrariness in any system set down. Limitations of resources – mental, computational, and statistical – enforce a model that, although complicated enough for our capacities, is yet enormously simple relative to the present state of understanding of the world we seek to explain. Until we can develop a simpler picture of the world, by an understanding of interrelations within sections of the economy, the construction of a model for the economy as a whole is bound to be almost a complete groping in the dark. The probability that such a process will yield a meaningful result seems to me almost negligible.

The model builders have, of course, recognized this problem. For example, it explains the distinction they make between a model – which is a class of admissible hypotheses – and a structure – which is a single hypothesis. It explains also their emphasis on examining the economic theory implicit in their equations, and on checking the signs of their statistically estimated parameters for 'reasonableness'.

In so far as they think the prospect more hopeful than I do, it is because they assess differently the existing state of our knowledge – they think we have more basis for narrowing the range of admissible hypotheses than I do. On this point, I venture to suggest that they have been misled by failing to distinguish among different kinds of economic theory. We do have a very well developed and, in my view, successful and useful theory of relative prices which tells us a great deal about relationships among different parts of our economic system, about the effects of changes in one part on its position relative to others, about the long-run effects of changes in technology, the resources at our disposal, and the wants of consumers. A theory of short-run changes in the economy as a whole must deal with many of the phenomena that are dealt with in price theory, and thus it is tempting to suppose that price theory substantially reduces the arbitrariness of a system of equations – enables us to narrow substantially the class of admissible hypotheses.

I believe that this is a serious mistake. Our theory of relative prices is almost entirely a static theory – a theory of position, not of movement. It abstracts very largely from just those dynamic phenomena that are our main concern in constructing a theory of economic change. The basic empirical hypothesis on which it rests is that the forces determining relative prices can be considered largely independent of the forces determining absolute prices; and its success is testimony to the validity of this hypothesis.

A theory of change cannot, of course, be constructed completely independently of a theory of relative prices. The two must in some sense be consistent with one another, and thus there is a real point in checking any theory of change to see that it does not have implications for the relations among the parts that are inconsistent with the theory of relative prices. The important point is that the existing theory of relative prices does not really help to narrow appreciably the range of admissible hypotheses about the *dynamic* forces at work.

Monetary theory, interpreted broadly, has somewhat more to offer. It is at least concerned with absolute prices. But even monetary theory, in its present state, is less useful than might at first appear. It too has typically been concerned with positions of equilibrium, with comparative statics rather than with dynamics – and this, I may add somewhat dogmatically, applies equally to Keynesian and pre-Keynesian monetary theories.

One cannot, of course, specify in advance what a workable theory of change will look like when it is developed. But I think it is clear that it will have to be concerned very largely with leads and lags, with intertemporal relations among phenomena, with the mechanism of transmission of impulses – precisely the kind of thing about which neither contemporary price theory nor contemporary monetary theory has much to say.

The direction of work that seems to me to offer most hope for laying a foundation for a workable theory of change is the analysis of parts of the economy in the hope that we can find bits of order here and there and gradually combine these bits into a systematic picture of the whole. In the language of the model builders, I believe our chief hope is to study the sections covered by individual structural equations separately and independently of the rest of the economy.

These remarks obviously have a rather direct bearing on the desultory skirmishing between what have loosely been designated the National Bureau and the Cowles Commission techniques of investigating business cycles. As in so many cases, the difference between the two approaches seems to me much greater in abstract discussions of method than it is likely to prove in actual work. The National Bureau has been laying primary emphasis on seeking to reduce the complexity of phenomena in order to lay a foundation for a theory of change; the Cowles Commission on constructing the theory of change. As the National Bureau succeeds in finding some order, some system, in the separate parts it has isolated for study its investigations will increasingly have to be concerned with combining the parts – putting together the structural equations. As the Cowles Commission finds that its general models for the economy as a whole are unsuccessful, its investigators will increasingly become concerned with studying the individual structural equations, with trying to find some order and system in component parts of the economy. Thus, I predict the actual work of the two groups of investigators will become more and more alike.

LAWRENCE R. KLEIN, *National Bureau of Economic Research*

Carl Christ has presented a splendid methodological account of a procedure for testing the validity of econometric models, but like many other econometric contributions of recent years it is weak in empirical or substantive content. I shall argue that his time series data contain an obvious gross error, that he has not chosen a desirable postwar revision of my prewar econometric model, and that his forecasting technique is both wrong and inefficient. Let me make matters quite clear at the outset, I do not accept any personal responsibility for anything that Christ has done. I participated to a negligible extent in his work.

The most serious deficiency in Christ's work is in the data he used for 1946-47 to revise my model and bring it up to date. These are critical observations since they provide the basis for revisions and in samples of 20-25 annual observations can play an important statistical role.[1] In addition, these data enter as lags in the forecasting for 1948. The series Christ has constructed show a drop in real aggregate output of more than 10 per cent from 1946 to 1947; this I do not believe. Every expert whose opinion I have canvassed concerning this period of experience advances the offhand guess that real output rose from 1946 to 1947. Christ shows an increase in total employment of 2.5 million persons from 1946 to 1947, and the Federal Reserve index of industrial production rose from 170 to 187 in the same period. Some of the trouble Christ finds with his estimates of the labor equation may well be traceable to these erroneous cross currents in employment and output. The same can be said about the estimates of the production function.

I am not prepared to give a full statement as to the source of his difficulties with the time series data, but it seems plausible to conclude that the price deflators used to pass from current dollar to constant dollar magnitudes are largely responsible. Price controls were lifted in the middle of 1946, and many of the published indexes rose to an excessive degree. I say *excessive* because many observers believe that official indexes seriously understated true prices toward the end of the war and in the early postwar period. The CIO argued vainly but correctly, I believe, that the BLS cost of living index was too low during the war. I shall not go into the reasons since this matter is discussed in other places.

Many of Christ's current dollar estimates could be wrong also because he used some questionable methods for converting the Department of Commerce national accounts from the new concepts (post July 1947) to the old concepts. His aim was apparently to reproduce an extension of my time series which followed the concepts I used as closely as possible. This, in itself, was a serious mistake. Looking at the supply-of-data situation of 1945 and 1946, I tried to do the best I could to get an adequate set

[1] Christ, in revising his paper, has written an unsatisfying appendix on the data problem. An interesting empirical finding of this appendix, however, is that hypothetical changes in the 1946 observations lead to radically different least-squares estimates of parameters of important equations.

of series covering the whole interwar period. I know only too well the deficiencies in my series. Looking at the supply-of-data situation of 1948 and 1949, if I were to have set out upon a revision of my model (Christ's situation), I would first have completely reworked all the series. The national accounts of the Department of Commerce, in their current state, would have been accepted as basic, and everything else would have been adjusted to them, including the pre-1929 data. Instead, Christ adopted the dubious procedure of forming regressions between the old and new concepts of the Department of Commerce during overlapping periods and extrapolated the data on the old concept to recent years from observations of the data on the new concept. Given Christ's questionable objective that he wanted to follow my outmoded time series as closely as possible, he could have done something more satisfactory. He could have accepted the most recent estimates of the Department of Commerce whenever no change of concept was involved, and he could have tried to estimate directly items that account for the change in concept, obtaining more satisfactory estimates of the series based on the old concepts.

In two other specific cases Christ used some data that seem obviously ill chosen. His price index of business capital goods (private producers' nonagricultural plant and equipment) is a weighted average of construction cost indexes and the wholesale price index for metals and metal products. The latter index is a poor substitute for an equipment price index in the postwar period. He then used this price index to extrapolate Fabricant's price index underlying business depreciation charges into the postwar years. It is almost certain that the depreciation deflator did not rise as fast as the price of newly purchased capital goods.

If we want to make a sound judgment about the use of econometric models for predicting some of the main economic magnitudes, we ought to reserve opinion until the most efficient use of the technique with available information has been tested. To forecast in the social sciences is difficult, and it is not likely that we shall get useful results with an inefficient application of any method. Christ's paper represents an inefficient application in many respects, and on the matter of data alone there are numerous things that he must do before he can draw any conclusions. The only really satisfactory approach open to him in the interests of efficiency is to revise all his series to agree with the new data of the Department of Commerce. This is not an easy task and will not appeal, of course, to 'sophisticated' econometricians who are more interested in less tedious research; however, it is imperative. Some day the 'sophisticated' econometrician will learn the trite result that his methods are not very useful when applied to poor data.

If this is too much of a job for Christ, a minimum task may be set out. He must first recalculate his estimates of the current dollar series without using a regression between the old and new concepts of the Department of Commerce. He should try to estimate directly the magnitudes that reconcile the two series. He should recompute all price deflators more carefully. If he still finds that real output fell from 1946 to 1947, he should revise the model on the basis of the 1947 estimate alone and recompute his extrapolations to 1948. I recommend this step because it seems likely that the main difficulty lies in the price deflators and that these are low in 1946 rather than high in 1947.

The only things for which I assume any responsibility are the construction of the prewar model and the forecasts, from it, for 1946 and 1947. My extrapolation to 1946 (*Econometrica,* April 1947, p. 134) estimated net national product in 1934 prices to be $121.6 billion. Christ's figure for the observed value is $115.2 billion. In terms of the customary accuracy involved in economic forecasts, this is not a bad correspondence. It is certainly in the right direction for the postwar situation. My forecast for *fiscal* 1947 (*ibid.,* p. 133) was $104.5 billion. Christ's figure for *calendar* 1947 is $103.3 billion, showing that my fiscal year forecast of real output was undoubtedly near the observed value. Since both my forecasts were made before the events occurred they had to use estimates of the relevant predetermined variables. Some of the estimates were not correct, but that, of course, is the case in any realistic forecasting situation.

The reason for introducing these considerations is to point out that Friedman's comments on Christ's paper cannot be accepted. Christ has not shown that econometric models break down as forecasting devices. We have two situations possible. Either Christ's series are accepted as correct, in which case I have been able to make some satisfactory forecasts from my model, or Christ's series are deemed incorrect, in which case Friedman cannot draw any substantive conclusions, as yet, from Christ's work. I am quite sure that the latter possibility is the correct one and that my forecast for 1947 was too low.

To many of us engaged in econometric work, it became obvious in the second half of 1947 that the most serious deficiencies in the existing models lay in the consumption equation and in the group of relations serving to determine absolute prices. During the war, households consumed at a low level in relation to their incomes; i.e., wartime observations on consumption and income lay substantially below the prewar consumption-income relationship. In the first postwar year (or 18 months), consumer spending worked its way back to the prewar relationship. In 1947 and 1948 the high levels of spending far surpassed the old relationship. We have, it appears, returned in 1949 to the neighborhood of the prewar relationship. The consequences of these movements are that the forecasts for 1946 were roughly correct and that those for 1947 and 1948 were low. Christ has tried to improve the consumption equation by introducing the real stock of cash balances as an additional variable. While there may be some plausibility to this approach, it is evident that it is not adequate. The observed consumption point for 1948 lies well above Christ's equation. In view of the low residual variation reported by Christ for the amended consumption equation in 1947, one might ask whether it would have been possible to know in advance of the 1948 forecast that this equation was going to give a low estimate for 1948. One thing apparent from the statistical estimates of the parameters is that the residual variation from Christ's consumption equation shows high serial correlation. Another shortcoming of his consumption equation is that it has a negative trend. I made two suggestions to Christ at the start of his work that would have improved the consumption equation and reduced the serial correlation of residuals. One change was to introduce lagged consumption as a separate variable in the consumption equation. Sound theoretical justification for this change can be developed. In the revised version of his paper, Christ made some least-squares estimates of a consumption equation of this type and found a remarkable improvement in the extrapolation to 1948. My other suggestion was to split consumption into categories such as durables, nondurables, and services, estimating separate equations for each. My own calculations of the equations for these components show a satisfactory lack of serial correlation in each set of residuals. If relative prices, stocks of consumer durables, and other relevant variables are taken into account, it is quite possible that a more satisfactory explanation of the postwar fluctuations in consumer spending could be obtained.

I find numerous faults in other equations of Christ's model and hold it unfortunate that such unreliable results should have been publicly presented when even Christ admits they contain some striking contradictions to common sense. I can think of many research possibilities that could be investigated to help clear up these faults and feel Christ should have investigated some of them before presenting his model. The production equation is unsatisfactory because a reliable or positive estimate of the marginal productivity of capital has not been found. The link between the prewar and postwar data on the stock of capital is very suspicious. Christ has, as I pointed out above, used a wrong deflator for depreciation in recent years; there is the problem of accounting for the transfer of surplus war property to private hands after the war. I have long insisted that the relevant variable for the production function is the flow of capital services rather than the existing stock of capital but find no real attempt on Christ's part to measure directly the flow of capital services. His indirect measurements, relating use of capital to net investment, strike me as being inadequate. These problems are of some importance, but two other defects of the production function seem more serious. In the first place, the man-hour concept of employment is far superior to the man-year concept used by Christ. He correctly notes the difficulty in preparing a series on man-hours going back to the early 1920's; however, I feel that some rough estimates should be made from the few sources available merely to see whether the bias in the employment data could have been responsible for the poor estimates of the marginal productivity of capital. My experience with American data shows that the trend influence (technological progress) in the production function has been very rapid, much more rapid than a simple linear function would allow. A quadratic trend would seem to fit an American model much better.

The production function finally selected (3.4) looks, as Christ has pointed out, suspiciously like his labor equation (4.2). An obvious alternative I would propose to avoid this difficulty is to express N as a function of $\frac{pX}{w}$, $\left(\frac{pX}{w}\right)_{-1}$ and t. This form is suggested by the theory of profit maximization subject to a Cobb-Douglas production function. Christ should consider this alternative. The other difficulties he encounters in his labor equation analysis can perhaps be explained by the inconsistent data he used on production and employment. I simply cannot believe the fantastic estimates he obtains for (4.0) since this was one of the most stable relationships of the interwar period, showing approximately the same structure for all methods of estimation in all models.

The limited-information estimates of the investment equation (1.0) seem equally implausible. Christ's equation is simply an extension of my own results, but I now lean towards a new formulation I recommended to Christ but which he did not try. I would express aggregate investment as a function of the current and lagged nonwage income originating in the sectors of the economy making the investments and the stock of capital. This is in accordance with the empirical findings in my paper (Part II). It is also possible that we may have written off too hastily the interest elasticity of investment as negligible. While I still do not think that investment for the whole economy is highly interest elastic, there is still some possibility of a small interest effect. I would favor an investment equation using nonwage income (profits before interest), the stock of capital, and bond yields as explanatory variables.

In essence, Christ's revision of the model has been to test the prewar equations against incorrect data of two postwar years, to recompute the parameters of the model with the two later observations, to rewrite the equations with a dubious production function[2] as one of the structural

[2] On the other hand, I find his wage equation (5.0), brought in by the revision, to be quite useful and satisfactory.

equations, and to use real cash balances as a variable in the consumption equation. As an alternative, I suggest a revision that I feel would be much more rewarding. First, rework the entire set of data as suggested in the first part of these comments. Secondly, try to use additional statistical information such as that provided in quarterly and cross-section data. Thirdly, revise the equations of the model as follows: The consumption, investment, production, and labor equations should be treated as suggested above. In case a satisfactory estimate of the production function is not obtained, the best alternative may be to rewrite the system in such a way that this equation does not appear explicitly but is imbedded in other equations. Industrial sectors and other components of variables should be treated in additional equations in a less aggregative model. New equations should be introduced to explain corporate savings and imports as endogenous variables. This plan of revision is not simple, but it is the direction nonsuperficial work must follow. Revisions like these will prove, I predict, to be much more valuable than any refinements of statistical methodology. I find Christ's empirical work disappointing in that it made practically no attempt to introduce more basic revisions such as these.

After testing and revising my model on the basis of estimated data for 1946 and 1947, Christ extrapolated the revised model to 1948, a year outside the sample observations. Although the mechanics of his extrapolating procedure are straightforward, I find his technique at serious fault from an econometric point of view.

An econometric model usually contains as many equations as there are endogenous variables thus enabling one to express the endogenous variables in terms of the predetermined variables, once the structural parameters are estimated. I find it very curious that Christ has gone to all the trouble of structural estimation and then has not used the estimated model to express endogenous magnitudes in terms of predetermined variables for purposes of extrapolation. Table 3 of Christ's paper contains his calculations underlying the predictive ability of the model. For the reader's benefit some comments are called for on this table. On the basis of a model involving 10 stochastic equations, Christ makes 13 predictions by one method and 21 by another. The structure of the model has been seriously violated, for it is not designed to yield more than 10 predictions. Christ goes through an elaborate procedure in testing and constructing a model of essentially 10 equations. He then throws away this information and makes 13 or 21 forecasts, in the latter case often getting more than one forecast for the same variable. The mechanics of this procedure are obvious, but the rationale is surely lacking. For example, equation 11.0 (C) in Table 2 gives us an estimate of a structural relation showing how the average interest rate depends on predetermined variables. This equation extrapolates tolerably well to 1948, predicting the interest rate to be 2.99 per cent; the observed value is 3.08 per cent. Christ estimates, in Table 2, the parameters of a structural equation showing how interest rates are related to predetermined variables, then turns to other equations to predict the interest rate in Table 3. The structural equation gives a much better prediction than the equations used in Table 3. It so happens that by throwing away information, Christ has biased his test of the predictive ability of the econometric model, since the structural equation extrapolates better than either of the naive models; whereas his reduced form predictions in Table 3 are worse than the naive model predictions.

Because of the faulty character of the data used and because of the inadequacy of the production and the demand-for-labor equations, I find it impossible to accept Christ's revised version of my model. However, I shall present an interesting experiment with the parts of the model that are more or less acceptable. Consider a system composed of CLS equations 1.0, 2.0, 6.5, 7.0, 10.0, 11.0, and definitions 12 and 13. In this model, equations related to the determination of rents in a free market setting are obviously suppressed because of rent controls, a fact Christ neglects for some unknown reason.

As stated previously, my former model was particularly weak in that actual consumption has been far above the consumption equation and there is no satisfactory scheme for the determination of absolute prices. I accept, for the moment, 6.5 (CLS) as the best possible version of the consumption equation until basic research in this area has progressed further, and take the price level as given until a suitable empirical scheme can be developed for incorporating this item into the model as an endogenous variable.[3] Least-squares estimates are used because it is the only type that has been calculated for the particular consumption equation used.

Solving this model for Y, C, I, D_1, D_2, i in terms of the other variables and substituting the observed values of the latter set in 1948 I find the following extrapolations:

$Y = \$68$ billion	$I = \$1.71$ billion	$D_2 = \$1.62$ billion
$C = \$73$ billion	$D_1 = \$1.90$ billion	$i = 2.99$ per cent

This model contains only 6 stochastic equations; hence there are only 6 extrapolations. These estimates are obviously defective as compared with observations, yet proceeding along the lines of Christ's paper we conclude that in 4 out of 6 cases they are better than those of either naive model.

One year's test tells us practically nothing in a statistical sense about the merits of econometric techniques. This is as true of my example as of

[3] This procedure is tantamount to a rejection of the basic revisions undertaken by Christ and an examination of a less adulterated version of my model.

Christ's paper. Absolutely no scientific conclusions can be drawn until many forecasts have been made under realistic forecasting conditions with efficient methods. My example certainly is not a demonstration of the usefulness of econometric model building; I merely offer it as a challenge to the acceptance of any substantive results in Christ's paper.

In addition to the fact that Christ uses the wrong equations for extrapolating the model beyond the sample points, the entire character of his prediction scheme is so mechanical that it loses much efficiency. It would be convenient if we had arrived at the final situation where forecasting could be reduced to purely mechanical operations, but we are only approaching such a situation, and there are many nonmechanical operations that any sensible forecaster would use together with the econometric model in its present form. For this reason, Christ's extrapolation cannot, in any sense, be considered as optimal, given his facilities. To be more specific, an econometric forecaster should be wary of structural change between the prewar and postwar period. Cross-section data from the Surveys of Consumer Finances and the surveys of investment intentions may throw substantial light on the postwar structure of the consumption and investment equations. Christ did not even consider this material. He could have used these data to check his estimates of consumption and investment or he could have, perhaps, used them to estimate the current structure of the consumption and investment equations to be used together with the other relationships of his model for extrapolation.

The assumption that there is no serial correlation of the disturbance terms of our econometric relationships may not be valid. Correction factors for the estimated values of the endogenous variables may be looked for in the trends in the most recent values of the estimated disturbances.

An alternative to taking into account the serial correlation in the disturbances would be to boost up some of the equations to make them conform with the postwar data as closely as with the prewar data. An objective way would be to introduce a dummy variable that takes on zero values in the prewar period and unit values in the postwar period.

It seems clear, in any case, that a competent forecaster would have used an econometric model on the eve of the prediction period far differently and more efficiently than Christ used his model.

In Appendix C Christ argues that the incorrectness of his data for the postwar years may have affected some of his structural equations but not the predictions made from the model. I find this argument weak in many respects. In the first place, the tests carried out for purposes of revising the model will be affected by a change in data if the estimated parameters are changed. Thus, at the earliest link of a chain of calculations there is weakness. This weak link spoils the entire chain. Secondly, the prediction equations used in the Appendix are the same as those used in Table 3. I reiterate that these are not the equations we want for purposes of forecasting. There is no easy solution from rough calculations like those of the Appendix. The only scientific way to approach the problem is to rework the entire set of basic data, obtain revisions of the model that stand up better than those offered by Christ (especially the consumption, production, and labor demand equations), and use the most efficient forecasting technique available. When all this has been properly done, we can come to the question of the predictive ability of econometric models.

REPLY BY MR. CHRIST

Lawrence Klein's comments have made it clear that my paper is not a finished piece of work containing well established results, and that its merits, such as they are, lie in the methodological field, illustrating the application of various methods of testing econometric models. I regret that I did not make this clear myself. I regard as the most valuable part of my paper the exposition and illustration of procedures for prediction and testing, rather than the particular model tested or the particular results arrived at, and that is the basis upon which I would like my work to be judged. Perhaps if I had made this clear, Klein would not have found it necessary to reiterate in his comments many of the criticisms and suggestions that already appear (some at his instance) in the paper and its appendices. To the extent that the paper does give the reader the impression that its results are reliable and should be accepted without further investigation, I think Klein is justified in many of his comments.

Klein's comments are divided into several parts, concerning (1) the data, (2) the form of the equations, (3) the number of variables predicted, (4) prediction from the linear reduced-form equations vs. prediction from the structural equations, (5) an experimental calculation of Klein's, (6) the question of mechanical vs. discretionary methods of prediction, and (7) my statement of the probable results of performing the desirable recomputations Klein suggests.

1) I have already essentially accepted most of Klein's comments on the data, as indicated in my Appendix C, with two exceptions worth noting: First, there is his statement that ". . . as to the source of [my] difficulties with the time series data, . . . it seems plausible to conclude that the price deflators used to pass from current dollar to constant dollar magnitudes are largely responsible." I have cited in Appendix C sources that lead me to disagree with this, and to believe that my regression procedure for extending the undeflated series to 1947 is equally responsible. Second,

there is the statement that he "simply cannot believe the fantastic [limited information] estimates" I obtain for the labor equation (4.0), and that these can perhaps be explained by inadequacies in the data. As I said in my paper, they appear fantastic to me as well, but I think it is clear, from the fact that my *least-squares* estimates of equation (4.0) are very reasonable and very close to Klein's least-squares *and* limited information estimates, that the data are not the controlling factor: if they were, my least-squares estimates would have been as absurd as my limited information estimates. I have commented on this matter in Section 12 of my paper.

2) Klein's comments on my choice of equations in the model are relevant to the question of how good the econometric technique is or can be for prediction purposes, because they suggest improvements that can be expected to lead to better results; they are not relevant to the process of testing the predictions of this particular model (though they may help to explain failures). I share most of Klein's criticisms on this point – in fact they are in my paper – except that I do not like to use a quadratic trend in the production function even though it fits the past data well: it is always possible to invent some function of time that fits a given set of data well or even perfectly, but where there are random elements such a function is not a reliable extrapolating device unless there is some substantive reason to believe that it will continue.

3) Klein states that my model "is not designed to yield more than 10 predictions" because it contains only 10 stochastic equations, and that I have made predictions of 13 variables. (Klein might be misunderstood when he says "Christ makes 13 predictions by one method and 21 by another"; the 21 predictions are actually predictions of the same 13 variables, with some duplication due to the possibility of using different sets of restrictions in estimating some of the parameters of the reduced form, as explained in Appendix E.) The three nonstochastic equations are the identities defining disposable income Y, private output X, and wage rate w (there is a fourth, defining capital stock K, but I ignored it because K appears as such nowhere else in the model). It is true that an arbitrary number of new variables could be added, each defined by a new identity, and that by a suitable choice of these new variables it would be possible to change the 'score', i.e., the number of variables predicted successfully by the model, from 6 out of 13 to, say, 16 out of 23, or even to 93 out of 100, without changing the original model in any way. This is what Klein means by saying that the number of variables that can legitimately be predicted by a model cannot exceed the number of stochastic equations in the model. However, there is a certain arbitrariness in deciding, regarding my model for example, which 3 variables should be eliminated by the 3 identities and which 10 should be predicted. Thus, if I had chosen to eliminate H, X, and N

respectively by the 3 identities, the score would have been raised from 6 successes out of 13 variables to 6 out of 10; on the other hand, if I had chosen to eliminate D_2, p, and either w or W_1, the score would have been lowered to 3 out of 10 vis-a-vis naive model I and to 4 out of 10 vis-a-vis naive model II. The identities as they are written suggest that the most natural variables to eliminate are Y, X, and w; had this been done, the score would have become 5 successes out of 10. Since each variable in the model has a real economic meaning and represents an interesting economic magnitude, it seems that there is arbitrariness involved in *eliminating* 3 by identities, as well as in *not* eliminating 3 and thereby having more variables to predict than stochastic equations.

4) Klein argues in favor of using structural equations instead of reduced form equations for making predictions (specifically, this means first to substitute known or assumed values of predetermined variables, and estimated values of parameters, into the structural equations; then to solve the resulting system of equations simultaneously for the values of the jointly dependent variables). In support he cites the interest rate, for which limited information estimates of the interest equation (11.0) give a better 1948 prediction than least-squares estimates of the reduced form. Now the interest rate is in a unique position in my model (and in Klein's) because it is already expressed in terms of predetermined variables in equation (11.0), which means that it can be predicted from the structural equations directly without the algebraic operations of simultaneous solution. If one seeks instead to predict price level p or disposable income Y from the structural equations, one finds that, because of the nonlinearities in the model (including the identities), simultaneous solution leads to a quintic equation in p or in Y, respectively. (Klein's model leads to a cubic in p or in Y, as he does not have the nonlinear identity defining wage rate.) I have not been able to obtain anything except absurd values for p and Y from calculations of this sort, for either my model or Klein's model, but I had not intended to refer to this until I could determine the reason. As matters stand now, it appears on common-sense grounds, as Klein says, that prediction from structural equations will usually be superior to prediction from estimates of linearized reduced form equations, because of using more restrictions, but the few cases I have tried do not bear this out. Further investigation of the reason is called for.

5) Klein has presented an experimental calculation, using a modified model and my data, to see what kind of predictions the structural equations make for 1948 if the model is not required to predict prices, the true 1948 price level being used as if it had been known when the predictions were made. The results are better than mine, as measured by my criterion of the number of cases in which the naive models are bested by the econo-

metric model (though the structural-equation predictions of the two most important variables, Y and C, are quite far off even compared with predictions from my least-squares estimates of the reduced form). Klein's forecasts for 1946 (*Econometrica,* April 1947), to which he refers in his comments, are likewise based on a model in which the price level is supplied from outside (by an enlightened guess) rather than predicted by the model. At this stage of econometric research it is apparently more accurate to guess at the price level, then use the econometric technique to predict other variables on the basis of that price level than to predict *all* variables by econometric methods.

6) This brings up an interesting issue. Klein's criticism of my forecasting methods as inefficient and mechanized springs from the fact that he and I are pursuing different objectives. His objective is to make good predictions *now* with the resources at hand. If this were my objective, I would not rely only on mechanical procedures any more than Klein would; I would try to use judgment in taking account of certain information outside the model, such as recent trends in calculated disturbances, recent cross-section studies, and if predictions were improved by guessing at the price level, I would guess at the price level as well as I could. But my objective is different: it is to make some progress toward evaluating a kind of prediction procedure that is scientific, in that as far as possible it is reproducible and free from discretionary judgments (of course this does not preclude the incorporation of cross-section data *into* the model). Such evaluation is needed because policy makers cannot be expected to have confidence in scientific forecasting techniques until such techniques are developed to the point where results are reproducible and independent workers can come to the same conclusions.

7) Klein says in his last paragraph that I argue in Appendix C that the difficulties in my postwar data "may have affected some of [my] structural equations but not the predictions made from the model". My words are: "the estimates of the [structural] parameters would be changed, and the 1948 fit of some structural equations would probably be improved, but there is no evidence that the predictions of important variables *by the reduced form* would be improved" (italics supplied). I agree with Klein's other comments in his last paragraph, apart from the question whether to use the structural equations or the reduced form for predicting, discussed under (4) above, and apart from the question of forecasting technique, discussed under (5) and (6) above.

GEOFFREY H. MOORE, *National Bureau of Economic Research*

It may be useful to put together the important results of Christ's Tables 2 and 3, and thereby bring out one or two additional conclusions (see accompanying table).

Column 3 shows the errors in 1948 predictions from Klein's least-squares estimates of his structural equations, which were based on data for 1920-41. Since the predictions are obtained by using actual 1948 values for the jointly determined variables, they are not strictly predictions. Nevertheless, the errors are large. If instead one had assumed that these variables would not change from one year to the next, the average absolute errors that would have been made since 1920 are those given in column 2; the errors that would have been made in 1948 are given in column 6. Comparing columns 3 and 2 we see that in the case of 7 of the 8 'predicted' variables the error from Klein's model is larger, often substantially larger, than the average annual absolute change in the variable. Comparison of columns 3 and 6 yields a similar result. We can infer, therefore, that the information extracted by the model from the *prewar* data was of very little use in 1948.

The errors in column 4 (Christ's model) are very much smaller than those in column 3. These too are not really forecasts, for the same reason. But the equations yielding these estimates use additional data for 1946 and 1947, whereas Klein used data ending in 1941; also, some changes were made in the actual form of the equations. The result is an improvement, though as judged by column 2 the improvement does not go very far. In the case of five variables the errors from Christ's model (col. 4) are smaller than the average annual changes (col. 2), and in five variables they are larger. The comparison with column 6, where the standard is the actual change 1947-48, is similar. Nevertheless, I think it is clear that Christ's model performed better than Klein's since every one of the errors is reduced substantially. One cannot help but feel, however, that if the addition of two years of data and a revision of some equations makes such a difference the models themselves must rest on a weak foundation. Christ's computations in Appendix C add to that impression. Here a change in the data for several variables for 1946 alone alters the 1948 predictions substantially.

A comparison between columns 4 and 5 is of some interest. Column 5 shows the results of the actual predictions from Christ's model. That is to say, the several variables were predicted from equations utilizing only the so-called predetermined variables. All except two of these predetermined variables are lagged, so that they would presumably be known at the end of 1947. On the other hand, as noted above, column 4 utilizes knowledge

Summary of Results of Tests

Variable	Actual value, 1948 (1)	Average annual absolute change in actual values, 1920-41, 1945-48 (2)	Observed minus predicted values, 1948: Using Equations Based on Data Ending 1941, and 1948 values of related variables (3)	1947, and 1948 values of related variables (4)	1947, and 1947 values of related variables (5)	Using Data for 1947 only, for the same variable (6)	1946-47 only, for the same variable (7)
Investment, bil. 1934 $	1.89	1.02	−2.93	−1.26	+0.89	−0.38	−1.33
Inventories, bil. 1934 $	34.3	1.25	−2.52	+1.18	+6.9	+3.0	+2.7
Price level, index, 1934:1.0	2.03	.07			−0.14	+0.15	−0.16
Production, bil. 1934 $	90.3	5.03		−8.90	+17.8	+3.0	+11.0
Employment, mil. man-years	52.1	1.81	+7.34	+2.14	+3.4	+1.1	−1.4
Wage bill, bil. current $	114.7	4.93			+7.1	+10.0	−3.2
Wage rate, thous. $ per man-year	2.20	.07		+0.17	+0.10	+0.15	−0.02
Consumption, bil. 1934 $	82.8	2.86	+12.26	+6.87	+2.6	+1.1	+5.9
Income, bil. 1934 $	79.9	4.15			+9.0	+1.7	+13.5
Owned housing, bil. 1934 $	1.86	.23	−3.07	−0.02	−0.18	+0.23	−0.06
Rent, index, 1934:1.0	1.24	.05	−0.044	+0.007	−0.01	+0.06	+0.03
Rental housing, bil. 1934 $	1.44	.22	−0.59	−0.14	−0.05	+0.28	+0.06
Interest rate, per cent	3.08	.38	−0.48	+0.07	+0.35	+0.22	+0.10

Derivation of columns:

(1) Col. 1 of Table 3.
(2) Col. 3 of Table 3.
(3) Col. 11 of Table 2, using Klein's least-squares estimates of structural equations (KLS).
(4) Col. 12 of Table 2, using Christ's least-squares estimates of structural equations (CLS).
(5) Col. 8 of Table 3, Christ's least-squares estimates of reduced form equations. The 1948 predictions are made from the 'predetermined variables', of which all except two are lagged 1947 values (see App. D).
(6) Col. 10 of Table 3 (Naive model I).
(7) Col. 11 of Table 3 (Naive model II).

of the actual 1948 values of relevant variables. It would seem then that the results in column 4 should be better than those in column 5 since not only is current information being used in column 4 but it is presumably more directly relevant to the predicted variable. Nevertheless, in only 5 of the 10 variables for which a comparison can be made are the errors in column 4 smaller than in column 5; in 4 cases they are larger; one (rent) is ambiguous because of the difference in the number of decimals. It seems then that the variables that were thought, when the model was constructed, to be directly relevant to the ones to be predicted are really not much more relevant than the so-called predetermined variables.

Christ has indicated the results of comparing columns 5 and 6 or 5 and 7. Another way of stating these results is to say that one could get as much information about the 1948 values for the several variables from their *own* values in 1947 or 1946-47 as one could from knowledge of the 1947 and earlier values of *other variables,* as used in his model. Put in this way, the results may not seem surprising, though it is clear that if the model were theoretically correct one would expect that a knowledge of the 1947 and earlier values of the variables other than the one being forecast would improve the forecast.

JAN TINBERGEN, *Netherlands School of Economics, Rotterdam*

1) In Section 3 one should, I think, add as a point in favor of the least-squares method that it requires a theory only for the equation studied and hence prevents the use of erroneous theories for the other equations.

2) It might have been of some use in Section 7 to mention as a possibility Modigliani's approach to the problem of the consumption function.

3) In Section 7 Christ states that he has added no new endogenous variables to the system by his modifications of the consumption function. I think this is true only as far as the statistical testing of his equations is concerned but not for the system looked at from the viewpoint of theory. In other words, the type of movements executed by the system will, of course, be influnced by M as an endogenous variable.